

IMPLEMENTASI METODE NAÏVE BAYES UNTUK MENDETEKSI STRES SISWA BERDASARKAN TWEET PADA SISTEM MONITORING STRES

Diva Fardiana Risa¹, Fajar Pradana^{*2}, Fitra Abdurrachman Bachtiar³

¹²³Universitas Brawijaya, Malang

Email: ¹divafardiana7@student.ub.ac.id, ²fajar.p@ub.ac.id, ³fitra.bachtiar@ub.ac.id

(Naskah masuk: 17 November 2020, diterima untuk diterbitkan: 17 November 2021)

Abstrak

Gangguan mental saat ini masih menjadi permasalahan bagi bidang kesehatan di seluruh dunia. Salah satu jenis dari gangguan mental yang dapat diprediksi saat ini adalah stres. Stres memiliki dampak yang sangat besar bagi Kesehatan penderitanya, namun masih banyak masyarakat yang terlalu meremehkan perihal keberadaan penyakit stres ini. Hal ini salah satunya disebabkan oleh media yang dapat digunakan untuk melakukan pengecekan tingkat stres masih sangat sedikit. Sejauh ini, pengecekan kondisi kesehatan mental khususnya stres dapat dilakukan melalui konsultasi ke psikolog terdekat. Namun, tidak banyak masyarakat yang mengetahui hal itu. Ketika seseorang mengalami gangguan kecemasan khususnya stres, maka ia akan cenderung melakukan tindakan yang dapat mengekspresikan kecemasannya di media sosial. Kegiatan ini dinamakan Self Disclosure. Hal ini dianggap dapat mengurangi beban penderita gangguan mental tersebut. Mengenai hal itu, saat ini penggunaan media sosial menjadi hal yang sangat lumrah dimasyarakat khususnya remaja. Salah satu jenis sosial media yang banyak digunakan oleh masyarakat adalah Twitter. Salah satu keunggulan Twitter adalah dikarenakan twitter lebih mudah digunakan dan memiliki tampilan yang sederhana. Selain itu, penulisan tweet pada akun twitter memiliki pembatasan jumlah karakter sehingga tweet yang ditulis pengguna lebih jelas dan ringkas. Oleh karena itu, pada penelitian ini akan dibangun fitur untuk mendeteksi tingkat stres melalui tweet pada akun twitter dengan menggunakan metode *Naïve Bayes* yang mana akan dapat mengklasifikasikan tingkat stres siswa berdasarkan tweet siswa kedalam tiga kelas yaitu kelas stres ringan, stres sedang dan stres berat. Fitur ini nantinya akan diimplementasikan pada sistem monitoring stres siswa berbasis website sebagai bahan pertimbangan bagi siswa dan guru bimbingan konseling dalam proses konseling siswa. Berdasarkan pengujian akurasi yang dilakukan dengan 90 data latih dan 4 data uji, maka didapatkan tingkat akurasi fitur ini mencapai angka 75%.

Kata kunci: Klasifikasi Teks, *Naïve Bayes*, Stres, Siswa, Twitter

IMPLEMENTATION OF THE NAÏVE BAYES METHOD FOR DETECTING STUDENT STRESS BASED ON TWEETS ON A STRESS MONITORING SYSTEM

Abstract

Mental disorders are currently still a problem for the health sector around the world. Stress is one type of mental disorder that can be predicted today. Stress has a huge impact on the health of sufferers, but there are still many people who underestimate the existence of the stress disease. This is partly due to the very few media that can be used to check stress levels. So far, checking mental health conditions, especially stress can be done through consultation with the nearest psychologist. However, not many people know about it. When a person experiences anxiety disorders, especially stress, he will tend to take actions that can express his anxiety on social media. This activity is called Self Disclosure. This is considered to reduce the burden on those with mental disorders. Regarding this, currently the use of social media is very common in society, especially teenagers. One type of social media that is widely used by the public is Twitter. One of the advantages of Twitter is that it is easier to use and has a simple interface. In addition, writing tweets on a Twitter account has a limit on the number of characters so that the tweet that the user writes is clearer and more concise. Therefore, this research will build a feature to detect stress levels via tweet on a twitter account using the Naïve Bayes method which will be able to classify students' stress levels based on student tweets into three classes, namely light stress, moderate stress and severe stress classes. This feature will later be implemented in a website-based student stress monitoring system as a consideration for students and counseling teachers in the student counseling process. Based on accuracy testing carried out with 90 training data and 4 test data, the accuracy rate of this feature is 75%.

Keywords: *Naïve Bayes, Stress, Student, Twitter, Text Classification*

1. PENDAHULUAN

Saat ini gangguan mental masih menjadi permasalahan di dunia kedokteran termasuk di Indonesia. Berdasarkan penelitian yang dilakukan, diketahui bahwa jumlah penderita gangguan kecemasan mencapai angka 9,8% dari keseluruhan 51.547 jumlah responden yang terdapat pada 21 negara, yang mana jenis penanganan dengan kualitas yang kurang baik banyak ditemukan pada negara berpenghasilan rendah (Alonso, et al., 2018). Menurut data yang dilaporkan oleh organisasi kesehatan dunia yaitu World Health Organization lebih dari 80% penyakit ini dialami orang-orang yang tinggal di negara yang berpenghasilan rendah dan menengah (WHO, 2017). Salah satu jenis dari gangguan mental yang dapat diprediksi saat ini adalah stres. Stres merupakan tahap awal dari runtutan penyakit gangguan mental seperti depresi, dan lain lain. Stres adalah kondisi akibat dari interaksi individu dengan lingkungan yang menimbulkan ketidaksesuaian antara tuntutan-tuntutan yang bersumber pada sistem biologis, psikologis, dan sosial individu (Sarafino, 2008). Stress juga memiliki tingkatan, yaitu stres tingkat ringan, stres tingkat sedang dan stres tingkat berat. Namun saat ini, masih banyak masyarakat Indonesia yang menganggap stres merupakan kondisi kesehatan yang tidak perlu penanganan. Padahal dampak dari stres yaitu selain dapat menghambat aktivitas manusia sehari hari juga dapat mengakibatkan depresi dan kecacatan fisik maupun mental.

Disamping itu, masyarakat Indonesia saat ini sangat gemar menggunakan media sosial. Saat ini, masyarakat cenderung mengungkapkan ekspresi dan perasaannya melalui sosial media karena dianggap lebih aman. Aktivitas ini dinamakan *Self Disclosure*. *Self Disclosure* dapat dilakukan dengan berbagai macam cara pada media sosial, seperti contohnya mengunggah foto, menulis status dan lain lain. Kegiatan *Self Disclosure* ini diketahui dapat mengurangi beban pengguna media sosial karena dianggap bahwa media sosial merupakan tempat yang aman dan menarik untuk mengungkapkan diri dan dapat menuai penghargaan sosial yang diinginkan (Forest & Wood, 2012). Pada penelitian yang dilakukan oleh Kominfo dan Lembaga UNICEF pada tahun 2014 mengenai perilaku anak dan remaja dalam menggunakan internet juga membuktikan bahwa 30 juta anak-anak dan remaja di Indonesia merupakan pengguna internet, dan media sosial menjadi pilihan utama dalam proses berkomunikasi. Salah satu media sosial yang sangat digemari oleh masyarakat Indonesia adalah Twitter. Pada laporan Twitter kuartal ke-3 tahun 2019 menyatakan bahwa pengguna aktif harian Twitter mengalami peningkatan sebanyak 17% dan mencapai angka 145 juta pengguna dan Indonesia merupakan salah satu negara yang memiliki partisipasi pengguna harian aktif dengan jumlah yang besar (Clinton, 2019).

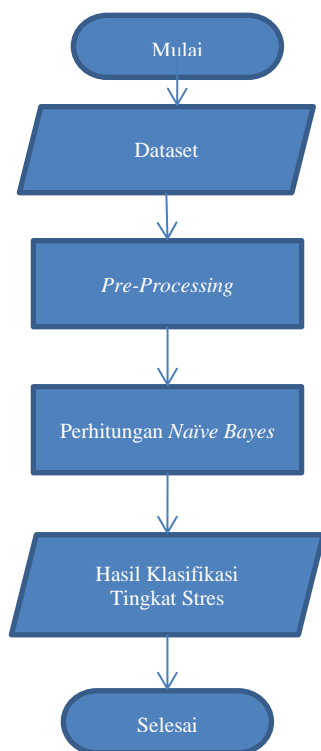
Proses pengelompokkan tweet pada tiga tingkatan stres akan dilakukan dengan menggunakan klasifikasi teks menggunakan teori *text mining*. *Text mining* merupakan proses menambang data dalam bentuk teks dimana sumber datanya didapatkan dari dokumen lain. Tujuan dari *text mining* adalah untuk mengekstrak informasi dari sumber data sehingga dapat digunakan untuk menganalisa dokumen uji (Putri & Setiadi, 2014). Dalam pemrosesan *text mining*, akan dilakukan dengan menggunakan metode *Naïve Bayes* yang dalam proses implementasi kedalam website sistem monitoring siswa akan menggunakan bantuan library Biobii. Pemilihan metode *Naïve Bayes* dilakukan karena metode ini memiliki tingkat akurasi dan kecepatan yang tinggi bila diterapkan pada jumlah data yang banyak dan memiliki perhitungan yang sederhana (Handayani & Pribadi, 2015).

Beberapa solusi telah ditawarkan dari penelitian sebelumnya salah satunya adalah mendeteksi tingkat stres menggunakan machine learning menggunakan sinyal electroencephalogram (EEG) dan metode *Naïve Bayes*. Penelitian ini bertujuan guna membagi kelas tingkat stres menjadi 4 level tingkatan, dari level 1 sampai level 4. Hasil penelitian ini menunjukkan bahwasanya menggunakan teknologi tersebut dapat mengklasifikasikan stres dengan nilai akurasi 94,6% (Subhani, et al., 2017).

Oleh karena itu berdasarkan latar belakang diatas, maka pada penelitian ini akan dibangun sebuah fitur untuk mendeteksi tingkat stres pengguna sistem berdasarkan status tweet pada akun twitter pengguna. Fitur ini akan diimplementasikan pada Website Sistem Monitoring Tingkat Stres Siswa yang ditujukan untuk Sekolah Tingkat Menengah Atas (SMA) untuk dapat dijadikan sebagai tambahan bahan pertimbangan bagi pengguna sistem yakni siswa dalam melakukan pengecekan tingkat stres dan hasilnya dapat dipantau oleh guru bimbingan konseling. Fitur mendeteksi stres ini memiliki cara kerja dengan mengambil data tweet terbaru siswa berdasarkan akun twitter siswa yang nantinya akan di proses menggunakan teori *text mining* dengan metode *Naïve Bayes* dan kemudian akan muncul hasil akhirnya berupa tingkatan stres siswa yang dibagi menjadi tiga kelas yaitu kelas stres ringan, stres sedang dan stres berat. Dalam proses pengambilan data tweet siswa akan dilakukan secara real time menggunakan bantuan Twitter API dan Library Thujhon. Sedangkan untuk implementasi metode *Naïve Bayes* pada *text mining* menggunakan bantuan Library Biobii. Oleh karena itu, diharapkan fitur pendeteksi stres siswa ini dapat meningkatkan kinerja website sistem monitoring stres siswa.

2. METODE PENELITIAN

Dalam melakukan penelitian ini, alur penelitian secara lengkap dapat dilihat pada Gambar 1 berikut.



Gambar 1. Alur Penelitian

Berdasarkan Gambar 1 diatas, maka dapat diketahui bahwa proses penelitian dimulai dari proses pengumpulan data yang digunakan, melakukan *pre-processing*, melakukan perhitungan *Naïve Bayes* dan menampilkan hasil klasifikasi tingkat stres.

2.1 Pengumpulan Data

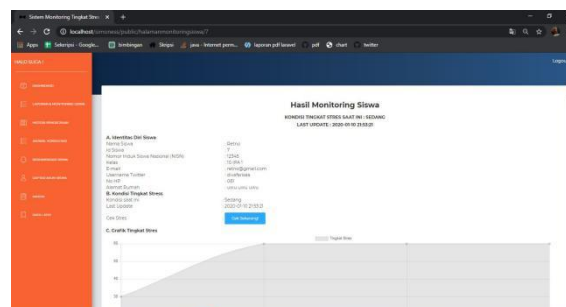
Pada penelitian ini proses pengumpulan data dilakukan dengan mengambil beberapa tweet secara acak pada akun twitter dengan menggunakan kata kunci tertentu yang berhubungan dengan stres dan kemudian data diklasifikasikan kedalam tiga tingkat stres yaitu stres tingkat rendah, sedang dan tinggi oleh pakar. Data ini adalah data yang digunakan sebagai data latih pada proses klasifikasi. Jumlah data latih yang digunakan adalah 90 data tweet dengan rincian 30 data tweet dengan kelas stres ringan, 30 data tweet dengan kelas stres sedang dan 30 data tweet kelas stres berat. Data latih ini akan disimpan dalam database sehingga dapat lebih mudah untuk digunakan. Contoh data tweet yang digunakan sebagai data latih tersedia dalam Tabel 1 berikut.

Tabel 1. Data Latih

D	Teks	Label
1	sayangku semangat hidupku terima kasih untuk semua yang engkau berikan aku cinta kamu	Ringan
2	biarlah semua terbakar agar semua benda ini melayang seperti debu di udara	Ringan
3	saat ini aku kecewa dengan keputusanmu yang aneh ini	Ringan
4	kelihatan sekali ada cewek lain mungkin dia lelah pantaslah sudah tidak ada waktu lagi	Sedang
5	terimakasih telah menemaniku selama tahun	Sedang

D	Teks	Label
	sampai jumpa di sana	
6	setelah berapa lama kamu telah terbiasa dengan hal hal gila kamu terbiasa untuk berjalan jalan dan merasa tidak berguna dan akhirnya kamu akan berhenti untuk memaafkan dirimu sendiri setiap orang yang kamu lihat kamu akan bandingkan dengan dirimu sendiri berteriak dan menangis di dalam diri	Sedang
7	oke aku pergi dari rumah penjara ini dan penghuni yang kayak kejam bilang kalau sudah tidak mau aku ikutin terus	Berat
8	semua ini tidak akan bisa aku lupakan tidak akan pernah bisa kecuali kalau aku mati	Berat
...
90	setelah aku mati aku akan menghantuimu siang dan malam aku tidak akan pernah meninggalkanmu lagi	Berat

Sedangkan untuk data uji yang merupakan tweet dari akun siswa diambil dari akun twitter dengan menggunakan bantuan Twitter API dan library Thujhon secara *realtime*. Data uji yang digunakan pada penelitian ini adalah 4. Setelah proses klasifikasi berhasil, maka hasil klasifikasi tingkat stres akan ditampilkan. Dimana, hasil klasifikasi tingkat stres ini juga nantinya akan disimpan di database sehingga dapat ditampilkan kembali kedalam website sistem monitoring tingkat stres dalam bentuk tabel dan grafik. Gambar 2 merupakan contoh antarmuka dari website sitem monitoring tingkat stres siswa.



Gambar 2. Antarmuka Halaman Hasil Monitoring Siswa Pada Sistem Monitoring Tingkat Stres Siswa

Gambar 2 diatas menampilkan halaman hasil monitoring siswa yang dapat diakses oleh siswa. Beberapa informasi yang tersedia pada halaman tersebut adalah diantaranya identitas siswa, grafik tingkat stres siswa dan lain lain.

2.2 Pre-Processing

Tahap *pre-processing* dilakukan dengan melakukan *tokenization*, *filtering* dan *stemming* seperti pada penjelasan sebelumnya. Tabel 2 menunjukkan data yang digunakan sebagai data uji, sedangkan Tabel 3 adalah hasil dari proses *pre-processing* yang dilakukan menggunakan 4 tweet yang digunakan sebagai data uji yang sebelumnya telah di klasifikasikan kelas stres nya oleh ahli.

Tabel 2. Data Uji

D	Teks	Label
1	Buat apa aku hidup cuma dibully lagian tidak ada yang peduli kalau aku mati aku cuma gadis misterius yang tidak pantas buat hidup	?
2	Aku kehilangan semuanya saat bersamaku yang ku dapat hanyalah merasa sakit yang dalam	?
3	Perasaan seperti keluarga bahagia Namun kini keadaan sudah berubah Kau pergi meninggalkanku dengan begitu mudah Tanpa memikirkan kami semua	?
4	Ya Allah ambillah nyawa ku buat apa aku ada kalau aku menyakiti hati orang lain	?

Setelah dilakukan *pre-processing* maka data uji menjadi seperti pada Tabel 3 berikut.

Tabel 3. Data Uji Setelah Di *pre-processing*

D	Hasil <i>pre-processing</i>
1	'buat', 'apa', 'hidup', 'cuma', 'bully', 'lagi', 'peduli', 'kalau', 'mati', 'cuma', 'gadis', 'misterius', 'pantes', 'hidup'
2	'hilang', 'semua', 'bersama', 'hanya', 'rasa', 'sakit'
3	'rasa', 'keluarga', 'bahagia', 'kini', 'keadaan', 'berubah', 'kau', 'pergi', 'ninggal', 'mudah', 'mikir', 'semua'
4	'Allah', 'ambil', 'nyawa', 'buat', 'apa', 'kalau', 'nyakiti', 'orang'

2.3 Menghitung Hasil Klasifikasi Teks Menggunakan Metode *Naïve Bayes*

Setelah mendapatkan hasil *pre-processing* dari dokumen uji dan dokumen latih maka selanjutnya adalah menghitung peluang masing masing term dengan masing masing label yakni kelas stres ringan sedang dan berat menggunakan perhitungan *Naïve Bayes*. Untuk menggunakan persamaan *Naïve Bayes* maka diharuskan menghitung $P(c)$ terlebih dahulu. Nilai $P(c)$ dihitung menggunakan Persamaan (2). Sehingga didapatkan apabila Nilai $P(c)$ dari masing masing kelas ringan, kelas sedang dan kelas berat yaitu pada Tabel 4 berikut.

Tabel 4. Hasil Perhitungan Probabilitas Awal Munculnya Kategori c

P(ringan)	P(sedang)	P(berat)
0,333	0,333	0,333

Selanjutnya untuk menghitung nilai dari $P(d/c)$ menggunakan Persamaan (3). Sehingga didapatkan nilai masing-masing $P(d/c)$ dari masing-masing dokumen uji terhadap masing-masing kelas stres yang terdapat pada Tabel 5 berikut.

Tabel 5. Hasil Perhitungan *Conditional Probability*

D	P(Dn Ringan)	P(Dn Sedang)	P(Dn Berat)
1	0,00000325959 9922	0,000000435180 3594	0,00000548261 2699
2	0,00055413198 67	0,000114887614 9	0,00007188314 428
3	0,00000007669 646874	0,000000006593 641809	0,00000004130 028399
4	0,00013853299 67	0,000114887614 9	0,00021564943 28

Kemudian selanjutnya akan dihitung nilai dari $P(c|d)$ untuk masing masing kelas stres dan masing masing dokumen uji dengan menggunakan Persamaan (1). Untuk nilai $P(d)$ boleh tidak dihitung karena memiliki nilai yang sama. Hal ini dikarenakan $P(d)$ masing masing kelas memiliki nilai yang sama dan pada perhitungan ini $P(d)$ bertindak sebagai penyebut. Tabel 6 menunjukkan hasil perhitungan $P(c|d)$ menggunakan Persamaan (1).

Tabel 6. Hasil Perhitungan Dengan Menggunakan Metode *Naïve Bayes*

D	P(ringan Dn)	P(sedang Dn)	P(berat Dn)
1	0,00000108544 6774	0,000000144915 0597	0,00000182571 0029
2	0,00018452595 16	0,000038257575 76	0,00002393708 704
3	0,00000002553 992409	0,000000002195 682723	0,00000001375 299457
4	0,00004613148 789	0,000038257575 76	0,00007181126 113

Berdasarkan Tabel 6 diatas, dapat diketahui nilai posterior dari masing masing kelas stres ringan, kelas stres sedang dan kelas stres berat terhadap masing masing data uji. Diketahui bahwa nilai *posterior* data uji 1 pada kelas stres ringan sebesar 0,000001085446774, *posterior* kelas stres sedang sebesar 0,0000001449150597 dan *posterior* kelas stres berat sebesar 0,000001825710029. Pada data uji 2, nilai *posterior* kelas stres ringan sebesar 0,0001845259516, *posterior* kelas stres sedang sebesar 0,00003825757576 dan *posterior* kelas stres berat sebesar 0,00002393708704. Pada data uji 3, nilai *posterior* kelas stres ringan sebesar 0,00000002553992409, *posterior* kelas stres sedang sebesar 0,000000002195682723 dan *posterior* kelas stres berat sebesar 0,00000001375299457. Selanjutnya untuk data terakhir yaitu data uji 4, nilai *posterior* kelas stres ringan sebesar 0,00004613148789, *posterior* kelas stres sedang sebesar 0,00003825757576 dan *posterior* kelas stres berat sebesar 0,00007181126113.

2.4 Hasil Klasifikasi Tingkat Stres

Tahap ini merupakan tahap untuk menentukan label kelas suatu teks ditentukan dengan memilih peluang yang terbesar. Untuk menentukan kelas dari masing-masing data ditentukan dengan memilih peluang kelas terbesar dari masing masing data.

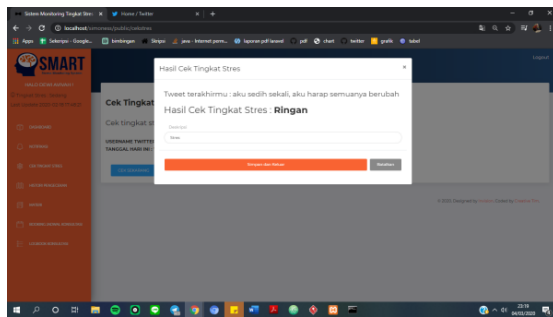
Sebagai contoh pada Tabel 4. Untuk data uji 1 yaitu "Buat apa aku hidup cuma dibully lagian tidak ada yang peduli kalau aku mati aku cuma gadis misterius yang tidak pantas buat hidup" memiliki nilai *posterior* kelas stres ringan sebesar 0,000001085446774, *posterior* kelas stres sedang sebesar 0,0000001449150597 dan *posterior* kelas stres berat sebesar 0,000001825710029. Apabila dibandingkan nilai hasil *posterior* antar kelas maka didapatkan hasil bahwa *posterior* kelas stres berat memiliki nilai yang terbesar dengan 0,000001825710029. Oleh karena itu, data uji 1

dikategorikan sebagai data uji kelas stres berat. Perbandingan ini juga berlaku untuk seluruh data uji.

Sehingga didapatkan kesimpulan bahwa berdasarkan perhitungan diketahui bahwa data 1 dikategorikan sebagai kelas stres tingkat berat, data 2 dikategorikan sebagai kelas stres tingkat ringan, data 3 dikategorikan sebagai kelas stres tingkat ringan dan data 4 dikategorikan sebagai kelas stres berat.

2.5 Implementasi Antarmuka Halaman Fitur Deteksi Stres Siswa

Tahap selanjutnya adalah mengimplementasikan antarmuka fitur deteksi stres siswa kedalam website sistem monitoring tingkat stres. Sehingga ketika diimplementasikan kedalam website sistem maka akan muncul seperti pada antarmuka sistem yang terdapat pada Gambar 3 berikut.



Gambar 3. Antarmuka Fitur Cek Tingkat Stres

Berdasarkan Gambar 3 yaitu antarmuka fitur cek tingkat stres, dapat diketahui terdapat beberapa informasi yang diperoleh yaitu diantaranya adalah menampilkan tweet terakhir yang diambil berdasarkan username pengguna yaitu siswa dan hasil cek tingkat stres yang didapatkan dari proses perhitungan *Naïve Bayes*, dimana hasil pengecekan ini memungkinkan tiga jenis stres yaitu stres ringan, stres sedang dan stres berat.

3. STRES

Stres merupakan suatu tekanan yang terasa menekan dalam diri individu, yang mana sesuatu tersebut dapat terjadi disebabkan oleh ketidakseimbangan antara harapan dan kenyataan yang diinginkan oleh individu baik keinginan yang bersifat jasmaniah dan rohaniah (Yusuf, 2004). Saat ini, stres sudah menjadi hal yang lumrah di masyarakat. Tingginya tuntutan hidup pada zaman yang juga terus menerus berkembang menjadikan banyak orang tidak dapat beradaptasi dengan tuntutan tersebut. Bahkan menurut Kupriyanov & Zhdanov (2014) menyatakan bahwa stres yang ada saat ini merupakan atribut kehidupan modern. Penyebab stres terdapat dua faktor yakni faktor dari diri sendiri maupun dari luar individu (Sukadiyanto, 2010). Yang banyak orang remehkan bahwa sebenarnya stres dapat mempengaruhi kondisi fisik seseorang

dikarenakan kondisi stres yang disebabkan oleh lingkungan sosial dapat mengaktifkan bagian otak yang menyebabkan rasa sakit secara fisik (Slavich, 2016). Tingkatan stres dibagi menjadi 3 yakni (Rasmund, 2004). yaitu stres tingkat ringan, stres tingkat sedang dan stres tingkat berat.

4. KLASIFIKASI TEKS MENGGUNAKAN METODE NAÏVE BAYES

Dalam penerapan teknologi pengklasifikasian teks, penulis menggunakan metode klasifikasi naïve bayes. *Naïve Bayes* classifier merupakan sebuah metode klasifikasi yang berakar pada teorema Bayes. Metode pengklasifikasian dengan menggunakan metode probabilitas dan statistik yaitu memprediksi peluang berdasarkan pengalaman di masa sebelumnya (*Teorema Bayes*) dengan ciri utamanya adalah asumsi yang sangat kuat (naif) akan ketergantungan dari masing-masing kondisi/kejadian (Rish, 2001).. Proses klasifikasi teks menggunakan metode *Naïve Bayes* memiliki beberapa tahap yaitu dimulai dengan tahap *pre-processing*, menghitung nilai probabilitas teks menggunakan metode *Naïve Bayes* dan kemudian hasil dari klasifikasi akan ditampilkan.

4.1 Pre-Processing

Langkah awal yang harus dilakukan adalah dengan melakukan *pre-processing* pada dokumen yang berupa teks. Langkah ini bertujuan untuk menyeragamkan keseluruhan dokumen untuk di proses selanjutnya agar lebih mudah. (Aji, et al., 2011). Dalam proses implementasinya, *pre-processing* memiliki beberapa tahap yaitu :

1. *Tokenizing* yaitu tahap pemisahan string input berdasarkan tiap kata yang disusun (Triawati, et al., 2009). Biasanya pemisah berupa spasi antar kata.
2. *Filtering* yaitu tahap untuk menghapus beberapa kata yang tidak diperlukan berdasarkan stop word list. Apabila kata tidak berada pada stopword list maka tidak perlu dihapus.
3. *Stemming* yaitu tahap untuk mentransformasikan kata yang terdapat di dokumen menjadi kata dasar atau *root word* nya dengan menggunakan aturan tertentu (Agusta, 2009). Biasanya, algoritma yang digunakan untuk mentransformasikan kata menjadi kata dasar dalam bahasa indonesia adalah Algoritma Sastrawi.

4.2 Perhitungan Naïve Bayes

Menghitung nilai probabilitas sebuah objek teks menggunakan *Naïve Bayes* adalah dengan menggunakan Persamaan (1) (Roonen, 2007).

$$P(c|d) = \frac{P(c) \times P(d|c)}{P(d)} \quad (1)$$

Keterangan :

$P(c|d)$ = *Posterior Probability* atau probabilitas kelas c diberikan dokumen d

$P(c)$ = *Prior Probability* atau probabilitas awal munculnya kategori c

$P(d|c)$ = *Likelihood* atau *conditional probability*

$P(d)$ = *Evidence* atau probabilitas munculnya dokumen d

Perhitungan *Prior* atau $P(c)$ pada *Naïve Bayes* dilakukan dengan menggunakan Persamaan (2).

$$P(c) = \frac{Nc}{N} \quad (2)$$

Keterangan :

Nc = Jumlah dokumen pada data latih yang masuk pada kategori c

N = Jumlah dokumen pada data latih

Perhitungan *Likelihood* atau $P(d|c)$ pada *Naïve Bayes* dilakukan dengan menggunakan Persamaan (3).

$$P(d|c) = P(w_1|c) \times P(w_2|c) \times \dots \times P(w_n|c) \quad (3)$$

Dimana,

$$P(w|c) = \frac{\text{count}(w,c)+1}{\text{count}(c)+|V|}$$

Keterangan :

$P(w_1|c)$ = *Condition probability* masing masing fitur

$\text{count}(w,c)$ = Jumlah kemunculan kata w pada kategori c

$\text{count}(c)$ = Jumlah total kemunculan semua kata pada kategori c

$|V|$ = Jumlah term unik atau fitur

5. MEDIA SOSIAL TWITTER

Media sosial adalah sebuah kelompok aplikasi berbasis internet yang dibangun diatas dasar ideologi dan teknologi Web, dan memungkinkan penciptaan dan pertukaran user-generated content (Kaplan & Haenlin, 2010). Media sosial mempermudah seseorang dalam proses berinteraksi dengan orang lain. Saat ini, media sosial sangat marak digunakan di dunia terutama di Indonesia. Twitter merupakan salah satu jenis media sosial yang ramai digunakan di Indoensia. Selain penggunaan yang mudah dan fitur yang beragam, twitter juga memberikan layanan API (*Application Programming Interface*) bagi para developer atau siapapun yang membutuhkan informasi dari Twitter. Di Indonesia, jumlah pengguna aktif Twitter mencapai 50% dari keseluruhan jumlah pengguna sosial media dan usia terbanyak dari pengguna sosial media adalah 17-34 tahun (Mardiyah, 2019).

6. HASIL DAN PEMBAHASAN

Selanjutnya setelah keseluruhan proses perhitungan telah selesai dilakukan maka hasil klasifikasi tingkat stres akan ditampilkan. Hasil dari perhitungan yang dilakukan oleh sistem dengan menggunakan metode *Naïve Bayes* menghasilkan kelas seperti yang terdapat pada Tabel 7 berikut.

Tabel 7. Hasil Klasifikasi Tingkat Stres Menggunakan Metode *Naïve Bayes*

D	Teks	Label
1	Buat apa aku hidup cuma dibully lagi tidak ada yang peduli kalau aku mati aku cuma gadis misterius yang tidak pantes buat hidup	Berat
2	Aku kehilangan semuanya saat bersamaku yang ku dapat hanyalah merasa sakit yang dalam	Ringan
3	Perasaan seperti keluarga bahagia Namun kini keadaan sudah berubah Kau pergi meninggalkanku dengan begitu mudah Tanpa memikirkan kami semua	Ringan
4	Ya Allah ambillah nyawa ku buat apa aku ada kalau aku menyakiti hati orang lain	Berat

Untuk melakukan pengujian ketepatan hasil perhitungan *Naïve Bayes*, maka pengujian akan dilakukan dengan menghitung tingkat akurasi dari hasil perhitungan sistem klasifikasi tingkat stres menggunakan metode *Naïve Bayes* dengan hasil klasifikasi tingkat stres oleh para ahli. Data uji sebelumnya telah diberikan kepada ahli untuk di tentukan kelas stres nya. Hasil klasifikasi tingkat stres berdasarkan tweet data uji oleh ahli terdapat pada Tabel 8.

Tabel 8. Hasil Klasifikasi Tingkat Stres oleh Ahli

D	Teks	Label
1	Buat apa aku hidup cuma dibully lagi tidak ada yang peduli kalau aku mati aku cuma gadis misterius yang tidak pantes buat hidup	Berat
2	Aku kehilangan semuanya saat bersamaku yang ku dapat hanyalah merasa sakit yang dalam	Sedang
3	Perasaan seperti keluarga bahagia Namun kini keadaan sudah berubah Kau pergi meninggalkanku dengan begitu mudah Tanpa memikirkan kami semua	Ringan
4	Ya Allah ambillah nyawa ku buat apa aku ada kalau aku menyakiti hati orang lain	Berat

Maka diketahui bahwa dari 4 data yang digunakan sebagai data uji diketahui bahwa hasilnya adalah 3 data benar dan 1 data salah. Data benar meliputi D1, D3 dan D4. Sedangkan D2 memiliki nilai salah.

Untuk menghitung tingkat akurasi metode. Maka menggunakan Persamaan (4) berikut.

$$\text{Accuracy} = \frac{n(\text{benar})}{n} \times 100\% \quad (4)$$

Keterangan :

Accuracy = Presentase tingkat akurasi

$n(\text{benar})$ = Jumlah dokumen yang terklasifikasi dengan benar

n = Jumlah dokumen keseluruhan

Sehingga apabila menggunakan persamaan untuk menghitung tingkat akurasi perhitungan metode, dapat disampaikan bahwa hasil klasifikasi tingkat stres menggunakan metode *Naïve Bayes* memiliki nilai akurasi 75%.

7. KESIMPULAN

Berdasarkan dari pembahasan sebelumnya, maka dapat ditarik kesimpulan diantaranya :

1. Penelitian ini membahas tentang pengklasifikasian tingkat stres siswa berdasarkan status tweet pada akun twitter menggunakan metode *Naïve Bayes* pada sistem monitoring tingkat stres berbasis website.
2. Tingkat stres akan diklasifikasikan menjadi 3 kelas yaitu kelas stres ringan, kelas stres sedang dan kelas stres berat. dan untuk memonitoring tingkat stres siswa.
3. Proses pengklasifikasian teks meliputi proses pengumpulan data yang terdiri dari data latih dan data uji, pre-processing, perhitungan *Naïve Bayes* kemudian hasil klasifikasi tingkat stres akan ditampilkan.
4. Dalam proses pengujian yang dilakukan menggunakan 90 data latih dan 4 data uji dan didapatkan bahwa tingkat akurasi dari klasifikasi tingkat stres menggunakan metode *Naïve Bayes* adalah sebesar 75%.

DAFTAR PUSTAKA

- AGUSTA, L., 2009. Perbandingan Algoritma Stemming Porter Dengan. Konferensi Nasional Sistem Dan Informatika, pp. 196-201.
- AJI, RIZQY BAYU., BAIZAL, ABDURRAHMAN, & FIRDAUS, YANUAR, 2011. Automatic Essay Grading System Menggunakan Metode Latent Semantic Analysis. Seminar Nasional Aplikasi Teknologi Informasi (SNATI), pp. 78-86.
- ALONSO, J., LIU, Z., LACKO, S. E. & SADIKOVA, E., 2018. Treatment Gap For Anxiety Disorders is Global : Results of the World Mental Health Surveys in 21 Countries. FOCUS ON: ANXIETY AND DEPRESSION WORLDWIDE, 35(3), pp. 185-282.
- CLINTEN, B., 2019. Kompas.com: Pengguna Aktif Harian Twitter Indonesia Diklaim Terbanyak. [Online] Tersedia di : <https://tekno.kompas.com/read/2019/10/30/16062477/pengguna-aktif-harian-twitter-indonesia-diklaim-terbanyak> [Diakses pada 12 Agustus 2020].
- FOREST, A. L. & WOOD, J. V., 2012. When Social Networking Is Not Working: Individuals With Low Self-Esteem Recognize but Do Not Reap the Benefits of Self-Disclosure on Facebook. PSYCHOLOGICAL SCIENCE, 23(3), pp. 295-302.
- HANDAYANI, F. & PRIBADI, F. D., 2015. Implementasi Algoritma *Naïve Bayes* Classifier. Jurnal Teknik Elektro , 7(1), pp. 19-24.
- KAPLAN, A. M. & HAENLIN, M., 2010. User of The World, Unite! The Challenges and Opportunities of Social media. Jakarta: Kompas Gramedia.
- KUPRIYANOV, R. & ZHDANOV., 2014. The Eustres Concept: Problems and Outlooks. JOUR, Volume 11, pp. 179-185.
- MARDIYAH, F., 2019. Websindo: Indonesia Digital 2019 : Media Sosial. [Online] Tersedia di : <https://websindo.com/indonesia-digital-2019-media-sosial/> [Diakses 12 Agustus 2020]
- PUTRI, E. & SETIADI, T., 2014. Penerapan Text Mining Pada Sistem Klasifikasi Email Spam Menggunakan *Naïve Bayes*. Jurnal Sarjana Teknik Informatika, 2(3), pp. 73-83.
- RASMUND, 2004. Stres, Koping dan Adaptasi. Jakarta: Sagung Seto.
- RISH, I., 2001. An Empirical Study of the Naïve Bayes Classifier. JCAI 2001 Work Empir Methods Artif Intell, Volume 3. pp. 41-46.
- ROONEN, F., 2007. The Text Mining Handbook: Advanced Approaches In Analyzing Unstructured Data.
- SARAFINO, E. P., 2008. Health Psychology : Biopsychosocial Interactions. 6th ed. New Jersey: USA: The College of New Jersey.
- SLAVICH, G. M., 2016. Life Stress and Health: A Review of Conceptual Issues and Recent Findings. Teaching of Psychology , 43(4), pp. 346-355.
- SUBHANI, A. R. ET AL., 2017. Machine Learning Framework for the Detection of Mental Stress at Multiple Levels. IEEE, Volume 5, pp. 13545-13556.
- SUKADIYANTO, 2010. Stres dan Cara Menanganinya. Cakrawala Pendidikan, Issue 1, pp. 55-64.
- TRIAWATI, C., BIJAKSANA M. A., INDRAWATI, N., & SAPUTRO W. A., 2009. Pemodelan Berbasis Konsep Untuk Kategorisasi Artikel Berita. Seminar Nasional Aplikasi Teknologi Informasi 2009, pp. 48-53.
- WHO, 2017. Depression And Other Common Mental Disorders. [online] Tersedia di : https://www.who.int/mental_health/management/depression/prevalence_global_health

[h_estimates/en/>](#) [Diakses pada 11 Agustus 2020].

YUSUF, 2004. Psikologi Perkembangan Anak dan Remaja. Bandung: Remaja Rosdakarya,