

Pemetaan Emosi Dominan pada Kalimat Majemuk Bahasa Indonesia Menggunakan *Multinomial Naïve Bayes*

(Mapping Dominant Emotion in Indonesian Compound Sentences Using *Multinomial Naïve Bayes*)

Wisnu Agastya¹, Aripin²

Abstract—This study aimed at mapping Indonesian sentences into emotion classes based on the classification process in those sentences. The results of emotion mapping can be applied in various fields, such as production of animated films and games, analysis of facial expressions, human-computer interactions, and development of other expressive virtual characters, specifically to produce facial expressions that match the spoken sentences. The method used for the emotion mapping process was the text classification using multinomial naïve Bayes model that was accompanied by dominant boundary equation. Multinomial naïve Bayes model in the text classification is used to determine the types and the emotional intensity of Indonesian sentences, whereas dominant boundary equation is used to determine the threshold in order to identify the dominant classes. The emotion classes used as references are six basic emotion classes according to Paul Ekman, i.e., happiness, sadness, anger, fear, disgust, and surprise. The experiment on the process of mapping emotions used Indonesian single and compound sentences. The experimental results show that the text classification using multinomial naïve Bayes model accompanied by dominant boundary equation can map compound sentences into several classes of dominant emotions.

Intisari—Makalah ini bertujuan untuk memetakan kalimat bahasa Indonesia ke dalam kelas-kelas emosi berdasarkan proses klasifikasi pada kalimat tersebut. Hasil pemetaan emosi dapat diaplikasikan dalam berbagai bidang, seperti produksi film animasi dan *game*, analisis ekspresi wajah, interaksi manusia-komputer, dan pengembangan karakter virtual ekspresif lainnya, khususnya untuk menghasilkan ekspresi wajah yang sesuai dengan kalimat yang diucapkan. Metode yang digunakan untuk proses pemetaan emosi adalah klasifikasi teks menggunakan model *Multinomial Naïve Bayes* yang disertai persamaan batas dominan. Model *Multinomial Naïve Bayes* pada klasifikasi teks digunakan untuk menentukan jenis dan intensitas emosi dari kalimat bahasa Indonesia, sedangkan persamaan batas dominan digunakan untuk menentukan ambang batas untuk mengidentifikasi kelas-kelas yang dominan. Kelas emosi yang digunakan sebagai acuan adalah enam kelas emosi dasar menurut Paul Ekman, yaitu senang, sedih, marah, takut, jijik, dan terkejut. Eksperimen pada proses pemetaan emosi menggunakan kalimat bahasa Indonesia yang meliputi kalimat tunggal dan kalimat

majemuk. Hasil eksperimen menunjukkan bahwa dengan klasifikasi teks menggunakan model *Multinomial Naïve Bayes* yang disertai persamaan batas dominan dapat memetakan kalimat majemuk ke dalam beberapa kelas emosi yang dominan.

Kata Kunci—Klasifikasi Emosi, Pemetaan Emosi, Kalimat Majemuk, *Multinomial Naïve Bayes*, Persamaan Batas Dominan.

I. PENDAHULUAN

Bahasa Indonesia, sebagai bahasa resmi dan bahasa nasional di Indonesia, berpotensi menjadi bahasa internasional. Hal ini dapat dilihat dari beberapa faktor yang dikemukakan oleh Badan Pengembangan dan Pembinaan Bahasa (BPPB), seperti jumlah penduduk Indonesia adalah yang terpadat ke-4 di dunia [1], peningkatan jumlah wisatawan asing [2], dan ekspatriat yang tinggal dan bekerja di Indonesia [3]. Bahasa Indonesia juga memiliki sistem bahasa dan ejaan yang mapan. Kemudian, penulisan pada bahasa Indonesia telah menggunakan huruf latin, sehingga dapat menjadi bahasa dalam berbagai bidang [4]. Telah dilakukan beberapa penelitian tentang bahasa Indonesia dalam bidang teknologi komputer, seperti *viseme*, pengolahan *text-to-audiovisual*, dan segmentasi kata pada bahasa Indonesia [5]-[9]. Hasil dari beberapa penelitian tersebut telah diterapkan pada pengembangan sistem pembelajaran bahasa Indonesia dan karakter virtual yang dapat berkomunikasi menggunakan bahasa Indonesia.

Struktur bahasa Indonesia terdiri atas unsur subjek, predikat, objek, pelengkap, dan keterangan. Secara garis besar, terdapat dua jenis kalimat bahasa Indonesia berdasarkan jumlah klausa/predikatnya, yaitu kalimat tunggal dan kalimat majemuk (majemuk, majemuk bertingkat, majemuk kompleks) [10]. Kalimat tunggal adalah kalimat yang mengandung satu predikat yang menyatakan hanya satu gagasan atau peristiwa, sedangkan kalimat majemuk memiliki lebih dari satu predikat yang menyatakan lebih dari satu gagasan dan umumnya mengandung konjungsi. Dalam sebuah percakapan atau komunikasi menggunakan bahasa Indonesia yang terjadi dalam kehidupan sehari-hari, penggunaan kalimat majemuk sering digunakan untuk mempersingkat beberapa gagasan.

Pada beberapa bidang, seperti produksi film animasi dan *game*, analisis ekspresi wajah, interaksi manusia-komputer, dan pengembangan karakter virtual ekspresif lainnya, komunikasi menggunakan bahasa Indonesia dituntut memiliki tingkat kealamian yang tinggi, sehingga dibutuhkan keterlibatan komunikasi nonverbal yang sesuai, khususnya ekspresi wajah. Dalam sebuah percakapan, kalimat bahasa Indonesia mengandung ekspresi emosi seseorang yang ditunjukkan pada

¹ Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Dian Nuswantoro, Jl. Imam Bonjol No. 207 Pendrikan Kidul Kota Semarang 50131 INDONESIA (tel: 024-3517261; fax: 024-3569684; e-mail: 111201609996@mhs.dinus.ac.id)

² Jurusan Teknik Biomedis Fakultas Teknik Universitas Dian Nuswantoro, Jl. Imam Bonjol No. 207 Pendrikan Kidul Kota Semarang 50131 INDONESIA (tel: 024-3517261; fax: 024-3569684; e-mail: arifin@dsn.dinus.ac.id)

ekspresi wajah pelaku percakapan. Untuk dapat menghasilkan ekspresi wajah yang sesuai dengan kalimat yang diucapkan pada karakter virtual, perlu dilakukan pemetaan emosi dari kalimat bahasa Indonesia.

Emosi adalah suatu reaksi terhadap seseorang atau kejadian [11]. Terdapat enam emosi dasar yang dikategorikan oleh Paul Ekman, yaitu senang, sedih, marah, takut, jijik, dan terkejut [12]. Emosi dapat dikenali melalui teks, audio (suara), dan visual. Pengenalan emosi melalui teks dipilih karena teks relatif lebih sederhana dan tidak berubah-ubah emosinya dibandingkan melalui visual atau audio [13]. Adapun emosi pada kalimat bahasa Indonesia dapat dikenali melalui kata sifat atau kata yang merepresentasikan emosi tersebut [14].

Pemetaan emosi pada kalimat bahasa Indonesia dilakukan menggunakan proses klasifikasi teks berdasarkan kelas emosi dasar. Adapun model algoritme yang dapat digunakan untuk proses klasifikasi teks antara lain, *k-NN*, *VSM*, *naïve Bayes*, dan *SVM*. Setiap model pada klasifikasi teks mampu mengenali jenis emosi dari sebuah teks [13], [15]-[18]. Namun, pada dasarnya setiap emosi memiliki intensitas yang memengaruhi pembentukan ekspresi wajah dan intonasi suara yang berbeda, sehingga dalam proses klasifikasi teks dibutuhkan model yang tidak hanya mampu mengenali jenis emosi, tetapi juga dapat menentukan intensitas emosi dari sebuah kalimat.

Multinomial naïve Bayes adalah pengembangan model *naïve Bayes* yang menghasilkan nilai probabilitas dari frekuensi atau jumlah kata yang muncul dalam sebuah kalimat. Frekuensi kemunculan kata dalam sebuah kalimat memengaruhi hasil dari nilai probabilitas tersebut. Hal inilah yang menjadi kelebihan model *multinomial naïve Bayes* untuk menentukan intensitas emosi dengan menghitung frekuensi kemunculan kata sifat dalam sebuah kalimat. Oleh karena itu, proses klasifikasi teks menggunakan model *multinomial naïve Bayes* dapat menghasilkan nilai probabilitas yang menentukan jenis dan intensitas emosi dari sebuah kalimat bahasa Indonesia. Namun, proses klasifikasi teks umumnya hanya menghasilkan satu kelas emosi saja, sedangkan kalimat majemuk yang memiliki lebih dari satu gagasan atau peristiwa berpotensi memiliki lebih dari satu emosi dasar yang berbeda.

Persamaan batas dominan adalah sebuah rumus yang diciptakan pada makalah ini untuk mengolah nilai probabilitas pada setiap kelas emosi yang dihasilkan dari model *multinomial naïve Bayes* menjadi beberapa jenis emosi dengan intensitas yang dominan. Metode persamaan batas dominan pada makalah ini ditujukan untuk mengatasi permasalahan klasifikasi teks pada pemetaan emosi kalimat majemuk yang hanya menghasilkan satu kelas emosi dasar. Persamaan batas dominan menghasilkan nilai batas dominan untuk menentukan emosi dominan pada setiap kalimat bahasa Indonesia. Pada dasarnya, setiap kalimat bahasa Indonesia mengandung seluruh unsur emosi dasar, tetapi intensitasnya berbeda-beda. Beberapa emosi dengan intensitas yang dominan memengaruhi informasi pada kalimat tersebut.

Terdapat beberapa penelitian yang terkait dengan permasalahan pada makalah ini. Salah satunya adalah penelitian mengenai sintesis ekspresi wajah berdasarkan teks

bahasa Indonesia menggunakan algoritme *naïve Bayes* dan *fuzzy logic* [13]. Dalam penelitian tersebut, model *naïve Bayes* berfungsi untuk menentukan emosi pada kalimat bahasa Indonesia berdasarkan lima kelas emosi (senang, sedih, marah, takut, jijik) dan metode *fuzzy logic* berfungsi untuk menentukan gabungan nilai dari lima kelas emosi, sehingga ekspresi wajah pada karakter 3D yang dihasilkan merupakan gabungan dari lima emosi dasar. Selanjutnya, penelitian tentang implementasi klasifikasi teks menggunakan model *naïve Bayes* menghasilkan respons emosi pada program *chatting* [15]. Dalam penelitian tersebut, dilakukan proses klasifikasi teks kalimat bahasa Indonesia berdasarkan tujuh kelas emosi (senang, takut, marah, sedih, jijik, malu, bersalah) menggunakan model *naïve Bayes* pada program *chatting* untuk menghasilkan emotikon yang sesuai dengan kalimat *input*. Adapun dalam penelitian tentang klasifikasi emosi pada teks bahasa Indonesia menggunakan algoritme *k-NN* [16] dilakukan proses klasifikasi teks menggunakan pembobotan TF-IDF dan model *k-NN* berdasarkan enam emosi dasar (senang, sedih, takut, marah, jijik, terkejut) yang menghasilkan nilai $k = 30-40$ sebagai nilai k yang optimal pada model *k-NN*. Penelitian lain tentang klasifikasi emosi pada teks bahasa Indonesia menggunakan algoritme *VSM* [17]. Dalam penelitian tersebut dilakukan proses klasifikasi teks kalimat bahasa Indonesia berdasarkan tiga kelas emosi (senang, takut, kecewa) menggunakan pembobotan TF-IDF dan model *VSM*. Ada juga penelitian tentang penentuan jenis emosi pada lirik lagu bahasa Indonesia menggunakan model *naïve Bayes* [18]. Dalam penelitian tersebut dilakukan proses klasifikasi teks berdasarkan empat kelas emosi (cinta, marah, sedih, senang) menggunakan model *naïve Bayes* untuk menentukan emosi dari setiap lirik lagu bahasa Indonesia.

Berdasarkan penelitian-penelitian tersebut, kontribusi dari makalah ini adalah memetakan lebih dari satu emosi dominan pada kalimat majemuk, yang umumnya pada proses klasifikasi teks hanya menghasilkan satu kelas sebagai ekspresi emosi pada setiap kalimat. Pada makalah ini, pemetaan emosi dilakukan pada kalimat bahasa Indonesia dengan struktur kalimat tunggal dan kalimat majemuk menggunakan klasifikasi teks menggunakan model *multinomial naïve Bayes* untuk mendapatkan jenis dan intensitas emosi. Sedangkan persamaan batas dominan digunakan untuk menentukan emosi dominan pada kalimat tunggal dan kalimat majemuk.

II. KAJIAN EMOSI DAN PSIKOLOGI

Emosi digambarkan sebagai keadaan perasaan yang berpengaruh pada perilaku, perubahan fisiologis dan pikiran. Emosi dapat dipengaruhi oleh rangsangan dari luar atau dalam diri seseorang. Emosi pada seseorang dapat menular pada orang yang memperhatikannya. Sebagai contoh, dua orang yang sedang terlibat emosi positif (senyuman, kegembiraan, kehangatan) akan membuat orang yang memperhatikan ikut merasakan bahagia, begitu pun dengan emosi negatif [14]. Emosi marah, sedih, senang, dan takut yang diungkapkan dengan kata sifat lebih banyak dibandingkan dengan emosi jijik, malu, dan terkejut [14]. Berikut adalah gambaran singkat dari enam emosi dasar pada manusia [12].

- Senang digambarkan sebagai keberhasilan atau kepuasan terhadap sesuatu.
- Sedih digambarkan sebagai kegagalan atau penyesalan.
- Marah digambarkan sebagai ganjalan atau frustrasi terhadap seseorang atau sesuatu.
- Takut digambarkan sebagai kekhawatiran terhadap suatu ancaman.
- Jijik digambarkan sebagai penolakan terhadap sesuatu dan rasa risi.
- Terkejut digambarkan sebagai perasaan terhadap sesuatu yang datang mendadak.

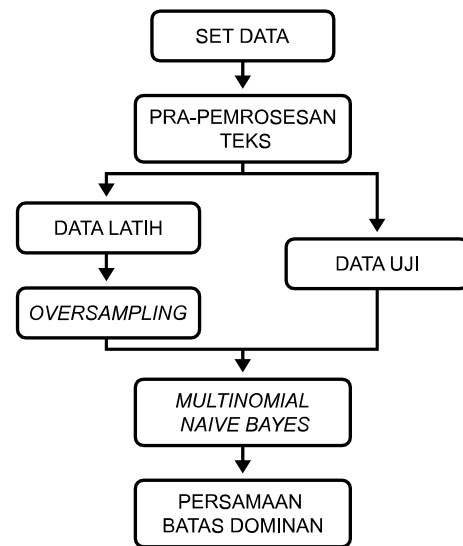
III. METODOLOGI

Tahapan yang dilakukan pada makalah ini terdiri atas enam langkah seperti yang ditunjukkan pada Gbr. 1. Secara garis besar, tahapan yang dilakukan dimulai dari pengumpulan data kalimat bahasa Indonesia menjadi set data. Kemudian set data diolah dalam tahap prapemrosesan teks untuk mengubah data tidak terstruktur (teks) menjadi data terstruktur. Selanjutnya, set data dibagi menjadi dua bagian, yaitu data latih dan data uji. Data latih berfungsi untuk proses *learning* atau pembelajaran model, sedangkan data uji berfungsi untuk proses pengujian model. Model yang digunakan pada proses klasifikasi teks adalah *multinomial naïve Bayes*. Hal ini bersangkutan dengan sifat model tersebut yang dapat mengolah kata dalam distribusi *multinomial*. Perbedaan frekuensi kemunculan kata pada setiap kalimat akan berbeda, sehingga pengolahan kata sifat menggunakan model *multinomial naïve Bayes* tidak hanya menghasilkan jenis emosi, tetapi juga intensitas dari emosi tersebut. Intensitas dari emosi tersebut digunakan untuk penentuan emosi yang dominan menggunakan metode persamaan batas dominan.

A. Akuisisi Data

Pada makalah ini, set data berisi kalimat tunggal dan kalimat majemuk yang dikumpulkan dari berbagai sumber informasi, seperti internet, buku, majalah, dan koran yang telah memenuhi unsur kelengkapan pola suku kata bahasa Indonesia atau fonem. Pemilihan kalimat tunggal bertujuan untuk mempermudah proses *learning* pada model *multinomial naïve Bayes*, sedangkan pada proses pengujian digunakan gabungan antara kalimat tunggal dan kalimat majemuk. Kemudian, ditentukan label atau kategori emosi pada set data berdasarkan kategori emosi dasar sesuai kata sifat dalam setiap kalimat. Dalam hal ini, penentuan kategori emosi dilakukan berdasarkan panduan ahli bidang psikologi. Hasil penentuan kategori emosi juga telah divalidasi oleh ahli bidang psikologi untuk memastikan setiap kalimat sesuai dengan kategori emosi yang telah ditentukan. Set data yang berhasil dikumpulkan berjumlah 1.143 kalimat yang telah dikategorikan berdasarkan kategori emosi seperti ditunjukkan pada Tabel I.

Namun, seperti yang ditunjukkan Tabel I, jumlah data pada setiap kelas emosi tidak seimbang. Masalah ketidakseimbangan jumlah data pada setiap kelas emosi dapat berdampak pada saat proses klasifikasi teks menggunakan model *multinomial naïve Bayes*, yaitu klasifikasi pada kalimat bersifat netral atau tidak mengandung ekspresi emosi akan mengambil kelas dengan



Gbr. 1 Diagram tahap penelitian.

TABEL I
JUMLAH ANGGOTA KELAS EMOSI

Kelas	Jumlah
Senang	276
Sedih	367
Marah	243
Takut	137
Jijik	91
Terkejut	29

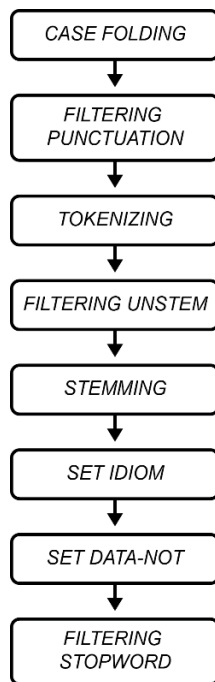
jumlah data yang paling sedikit atau disebut kelas minoritas. Masalah ketidakseimbangan jumlah data pada kelas emosi terjadi karena pada bahasa Indonesia, emosi marah, sedih, senang, dan takut yang diungkapkan dengan kata-kata sifat lebih banyak dibandingkan dengan emosi jijik, malu, dan terkejut [14]. Oleh karena itu, untuk mengatasi masalah pada ketidakseimbangan jumlah data pada setiap kelas, perlu dilakukan proses *oversampling* untuk mendapatkan jumlah data yang seimbang pada setiap kelas emosi.

B. Prapemrosesan Teks

Pada makalah ini, prapemrosesan teks dilakukan terhadap set data untuk mengubah data tidak terstruktur (data teks) menjadi data terstruktur. Tahap prapemrosesan teks dilakukan dalam delapan langkah yang ditunjukkan pada Gbr. 2. Tahap *filtering stopwords* diletakkan pada tahap akhir, dengan alasan terdapat tiga tahap tambahan yang membutuhkan beberapa kata *stopword*, seperti *filtering unstem*, *set idiom*, dan *set data-not*.

1) *Case Folding*: Tahap ini merupakan proses mengubah seluruh huruf dalam data teks menjadi huruf kecil. Proses ini berfungsi untuk mengurangi perbedaan penulisan huruf besar atau kecil pada seluruh set data.

2) *Filtering Punctuation*: Tahap ini merupakan tahap menghilangkan tanda baca dan nomor pada set data. Proses ini dilakukan karena nomor dan tanda baca dianggap tidak memiliki informasi yang berkaitan dengan kelas yang dipilih.



Gbr. 2 Diagram tahap prapemrosesan teks.

3) *Tokenizing*: Tahap ini adalah tahap pemisahan setiap kata pada setiap teks. Ini dilakukan untuk mempermudah proses perhitungan model algoritme.

4) *Filtering Unstem*: Pada tahap ini dilakukan *filtering* terhadap kata yang tidak boleh masuk dalam proses *stemming*. Kata tersebut akan berbeda maksud apabila diubah menjadi kata dasar, misalnya “meninggal” menjadi “tinggal”. Maka, dilakukan proses *filtering unstem* terhadap kata tersebut menggunakan daftar kata *unstem* yang telah dibuat.

5) *Stemming*: Pada tahap ini dilakukan proses *stemming* menggunakan *stemmer* Sastrawi [19]. Proses ini mengubah setiap kata pada set data menjadi kata dasar dengan menghilangkan imbuhan. Sebagai contoh, kata “melatih” diubah menjadi “latih”. Proses ini berfungsi untuk meningkatkan akurasi algoritme klasifikasi yang digunakan.

6) *Set Idiom*: Pada tahap ini dilakukan proses menggabungkan kata yang masuk dalam daftar kata idiom, contohnya “tanggung jawab” dan “jatuh cinta”. Kata tersebut digabungkan secara otomatis menggunakan daftar kata idiom yang telah dibuat. Proses ini berfungsi untuk menghindari kata idiom yang dipisah, karena kata tersebut akan berbeda maksud apabila berdiri sendiri.

7) *Set Data-Not*: Pada tahap ini dilakukan proses menggabungkan kata pengecualian dengan kata yang ada setelah kata tersebut menggunakan daftar kata pengecualian. Sebagai contoh, “tidak suka” akan berbeda maksud dengan “suka”. Maka, proses ini berfungsi untuk mempertahankan maksud dari setiap kata pada set data.

8) *Filtering Stopword*: Tahap ini melakukan proses menghilangkan kata sambung menggunakan daftar kata *stopword*, seperti “yang”, “di”, dan “pada”. Proses ini berguna untuk

TABEL II
JUMLAH ANGGOTA KELAS EMOSI SETELAH *OVERSAMPLING*

Kelas	Jumlah
Senang	365
Sedih	367
Marah	365
Takut	362
Jijik	364
Terkejut	364

menghindari kata-kata tersebut masuk dalam proses perhitungan klasifikasi. Proses *filtering stopwords* dilakukan pada akhir tahap prapemrosesan teks untuk menghilangkan sisa-sisa kata *stopword* dari tahap *filter unstem*, *set idiom*, dan *set data-not*. Setelah tahap ini, set data dipisahkan dalam tahap pembagian data.

C. Oversampling

Oversampling adalah tahap penggandaan anggota dalam kelas minoritas guna mendapatkan jumlah data kelas yang seimbang. Proses ini bertujuan untuk mengatasi masalah ketidakseimbangan jumlah data pada setiap kelas emosi. Pada makalah ini, metode *oversampling* hanya dilakukan pada data latih, karena proses penentuan ekspresi emosi ditentukan oleh data latih. Metode *oversampling* umumnya diikuti juga dengan metode *undersampling*, tetapi pada makalah ini hanya digunakan metode *oversampling*, karena metode *undersampling* dapat mengurangi jumlah data pada kelas mayoritas. Hal ini dapat memengaruhi berkurangnya jumlah data uji pada proses *learning*, yang dapat mengurangi akurasi pada model *multinomial naïve Bayes*. Hasil dari tahap *oversampling* ditunjukkan pada Tabel II. Terlihat bahwa setiap kelas emosi telah memiliki jumlah data yang seimbang. Kemudian, tahap selanjutnya adalah proses klasifikasi teks menggunakan model *multinomial naïve Bayes*, yang menghasilkan nilai probabilitas pada setiap kelas emosi.

D. Multinomial Naïve Bayes

Pada tahap klasifikasi teks, digunakan model *multinomial naïve Bayes* untuk mendapatkan jenis dan intensitas emosi. *Multinomial naïve Bayes* menghasilkan nilai probabilitas dari frekuensi atau jumlah kata yang muncul pada setiap kalimat. Frekuensi kemunculan kata dibutuhkan untuk menentukan intensitas emosi dari setiap kalimat. Intensitas emosi pada setiap kalimat yang berbeda berpengaruh pada saat penentuan emosi dominan menggunakan metode persamaan batas dominan.

Namun, pada model *multinomial naïve Bayes* terdapat masalah apabila sebuah kata tidak terdapat dalam kelas mana pun. Hal ini dapat berdampak pada terjadinya nilai probabilitas 0 atau *zero probabilities*. Maka, untuk menghindari *zero probabilities*, pada makalah ini digunakan metode *Laplace smoothing*. Metode ini bekerja dengan menambahkan nilai α , dengan syarat nilai $\alpha > 0$. Nilai α secara *default* adalah 1 [20].

$$P(w_i|c_j) = \frac{\text{count}(w_i,c_j) + \alpha}{\text{count}(c_j) + |\mathcal{V}|} \quad (1)$$

dengan $P(w_i/c_j)$ adalah nilai probabilitas kata i terhadap kelas j , $count(w_i,c_j)$ adalah nilai kemunculan kata i pada kelas j , dan α adalah nilai *Laplace smoothing* ($\alpha = 1$). Kemudian, $count(c_j)$ adalah jumlah anggota kelas j dan $|V|$ adalah jumlah anggota seluruh kelas tanpa pengandaan [21].

Algoritme *multinomial naïve Bayes* adalah model pada proses klasifikasi teks yang mengambil nilai probabilitas tertinggi untuk menentukan kelas emosi dari sebuah kalimat. Namun, untuk mengatasi klasifikasi teks pada kalimat majemuk yang berpotensi memiliki lebih dari satu kelas emosi yang berbeda, ditambahkan proses penentuan kelas emosi dominan menggunakan metode persemaan batas dominan.

E. Penentuan Emosi Dominan

Pada tahap ini, nilai probabilitas dari setiap kelas emosi dikonversi menjadi nilai persentase menggunakan (2), untuk mendapatkan nilai yang universal, sehingga dapat ditentukan nilai batas dominannya. Hal ini dilakukan karena total nilai probabilitas yang dihasilkan model *multinomial naïve Bayes* dapat berbeda jumlahnya pada setiap kalimat tergantung panjang kalimat tersebut.

$$Prosentase_i = \frac{Probabilitas_i}{\sum Probabilitas} \times 100 \% \quad (2)$$

dengan $Prosentase_i$ adalah nilai persentase dari emosi ke- i , $Probabilitas_i$ adalah nilai probabilitas dari emosi ke- i , dan $\sum Probabilitas$ adalah total nilai probabilitas dari seluruh kelas emosi.

Setelah nilai persentase dari setiap kelas emosi diperoleh, dapat ditentukan nilai batas dominan menggunakan (3) untuk menentukan kelas emosi yang memiliki nilai dominan dengan syarat nilai persentase emosi \geq nilai batas dominan.

$$dominan = min(p) + \beta \quad (3)$$

dengan $dominan$ adalah nilai batas dominan untuk menentukan kelas dominan, $min(p)$ adalah nilai persentase terkecil pada kalimat tersebut, dan β adalah nilai tambahan yang berfungsi membedakan kalimat netral dan kalimat yang memiliki emosi dominan. Proses menentukan nilai β akan ditunjukkan pada bagian hasil eksperimen.

IV. HASIL EKSPERIMEN

Pada hasil eksperimen ini ditunjukkan evaluasi model algoritme *multinomial naïve Bayes* pada klasifikasi teks menggunakan *confusion matrix* untuk mengetahui keakuratan model *multinomial naïve Bayes* dalam mengenali emosi pada kalimat bahasa Indonesia. Kemudian, untuk menunjukkan optimasi metode persamaan batas dominan untuk mengenali emosi pada kalimat majemuk, ditunjukkan perbandingan hasil pemetaan emosi tanpa metode persamaan batas dominan dan pemetaan emosi menggunakan persamaan batas dominan pada kalimat tunggal dan kalimat majemuk. Dalam hal ini, pemetaan emosi tanpa persamaan batas dominan hanya mengandalkan proses klasifikasi teks menggunakan model *multinomial naïve Bayes*, sedangkan pemetaan emosi menggunakan persamaan batas dominan adalah proses klasifikasi teks menggunakan model *multinomial naïve Bayes* yang disertai persamaan batas dominan.

TABEL III
KINERJA MODEL *MULTINOMIAL NAÏVE BAYES*

Data Latih	Data Uji	Precision	Recall	F1-Score	Accuracy
10%	90%	52,70%	52,70%	52,70%	52,70%
20%	80%	61,51%	58,40%	58,23%	58,40%
30%	70%	62,51%	61,29%	60,86%	61,29%
40%	60%	63,04%	61,91%	61,57%	61,76%
50%	50%	66,56%	65,79%	65,59%	64,24%
60%	60%	69,52%	67,29%	66,57%	67,29%
70%	70%	70,94%	69,28%	69,50%	69,28%
80%	20%	73,98%	73,58%	73,62%	73,58%
90%	10%	77,44%	75,47%	75,43%	75,47%

TABEL IV
KALIMAT UJI

No.	Kalimat	Jenis Kalimat
1.	kurang ajar siapa yang berani melakukan ini, apa jangan-jangan itu hantu ya	majemuk
2.	aku lega dia tidak mengetahuinya, namun aku khawatir dengan resiko yang akan datang	majemuk
3.	tempat itu memang horor, tetapi aku harus mengambil tasku yang tertinggal	majemuk
4.	aku mendapat kado dari ayah, namun betapa terkejutnya aku ternyata ini adalah sepatu yang selama ini aku inginkan	majemuk
5.	aku sedih berat badanku naik, aku kesal dibully terus sama teman-temanku	majemuk
6.	aku ingin menghajar para gerombolan itu, tapi nyaliku ciut aku hanya sendirian	majemuk
7.	akhirnya kegiatan jelajah malam kali ini selesai, namun ketika aku hendak ke toilet aku dikejutkan ular yang tiba-tiba di depanku akupun lari terburit birit	majemuk
8.	Aku tidak menyangka dia akhirnya pulang, aku memeluknya sambil menangis	majemuk
9.	kasihan sejak usia delapan tahun andi telah ditinggalkan ayahnya tetapi sekarang dia sudah sukses dan bahagia	majemuk
10.	malam ini malam jumat kliwon cintya tidak berani tidur sendirian	tunggal
11.	Presiden Jokowi hari ini melantik menteri baru bernama Handoko Budianto	netral

A. Evaluasi Model

Tabel III menunjukkan evaluasi model *multinomial naïve Bayes* pada klasifikasi teks menggunakan *confusion matrix*. Pada tabel tersebut ditunjukkan nilai *precision*, *recall*, *f1-score*, dan *accuracy* dengan sembilan kali percobaan menggunakan data latih = 10% dan data uji = 90% sampai dengan data latih = 90% dan data uji = 10%. Nilai tertinggi dihasilkan pada data latih = 90% dan data uji = 10%, yang memberikan nilai *precision* = 77,44%, *recall* = 75,47%, *f1-score* = 75,43%, dan *accuracy* = 75,47%.

B. Pemetaan Emosi

Pada Tabel IV ditunjukkan sebelas kalimat uji yang berisi satu kalimat tunggal, sembilan kalimat majemuk, dan satu

TABEL V
NILAI PROBABILITAS KALIMAT UJI

No.	Senang	Sedih	Marah	Takut	Jijik	Terkejut
1.	0,0118753	0,03582	0,54428	0,22563	0,01217	0,170213
2.	0,1440922	0,01521	0,00764	0,81764	0,00770	0,007707
3.	0,0499116	0,10026	0,05035	0,69788	0,05079	0,050795
4.	0,5050948	0,12737	0,01076	0,00527	0,00540	0,346100
5.	0,0170309	0,58331	0,34771	0,01703	0,01745	0,017457
6.	0,0283492	0,02834	0,43151	0,45408	0,02885	0,028851
7.	0,0008589	0,00087	0,00356	0,17895	0,00451	0,811238
8.	0,0283563	0,51192	0,02860	0,02835	0,02885	0,373902
9.	0,3229302	0,56783	0,02730	0,02687	0,02752	0,027525
10.	0,0006828	0,00068	0,00139	0,99583	0,00070	0,000700
11.	0,1646746	0,16600	0,16733	0,16467	0,16865	0,168659

TABEL VI
NILAI PERSENTASE KALIMAT UJI

No.	Senang	Sedih	Marah	Takut	Jijik	Terkejut
1.	1,19%	3,58%	54,43%	22,56%	1,22%	17,02%
2.	14,41%	1,52%	0,76%	81,76%	0,77%	0,77%
3.	4,99%	10,03%	5,04%	69,79%	5,08%	5,08%
4.	50,51%	12,74%	1,08%	0,53%	0,54%	34,61%
5.	1,70%	58,33%	34,77%	1,70%	1,75%	1,75%
6.	2,83%	2,83%	43,15%	45,41%	2,89%	2,89%
7.	0,09%	0,09%	0,36%	17,90%	0,45%	81,12%
8.	2,84%	51,19%	2,86%	2,84%	2,89%	37,39%
9.	32,29%	56,78%	2,73%	2,69%	2,75%	2,75%
10.	0,07%	0,07%	0,14%	99,58%	0,07%	0,07%
11.	16,47%	16,60%	16,73%	16,47%	16,87%	16,87%

kalimat bersifat netral. Kemudian dilakukan proses pemetaan emosi menggunakan model *multinomial naïve Bayes* untuk menghasilkan nilai probabilitas dari setiap emosi pada kalimat uji yang ditunjukkan pada Tabel V. Terlihat pada Tabel V bahwa nilai probabilitas yang dihasilkan dari model algoritme *multinomial naïve Bayes* tidak memiliki acuan pasti, dalam hal ini berarti total nilai probabilitas yang dihasilkan setiap kalimat memiliki panjang yang berbeda-beda, bergantung pada jumlah kata dalam kalimat tersebut. Oleh karena itu, dibuat nilai persentase yang merupakan hasil konversi dari nilai probabilitas untuk dapat menentukan nilai batas dominan. Hasil nilai persentase pada kalimat uji ditunjukkan pada Tabel VI.

Setelah nilai probabilitas diubah menjadi nilai persentase seperti yang ditunjukkan pada Tabel VI, dapat dilihat bahwa penggunaan model algoritme *multinomial naïve Bayes* dapat menghasilkan lebih dari satu kelas emosi dengan nilai dominan pada kalimat majemuk dan hanya menghasilkan satu emosi yang dominan pada kalimat tunggal. Sementara itu, pada kalimat bersifat netral, setiap kelas emosi menghasilkan nilai persentase yang hampir sama. Namun, untuk membuat mesin (komputer) dapat menentukan emosi dengan nilai yang dominan, metode persamaan batas dominan berperan

TABEL VII
KALIMAT PENGUJIAN PERSAMAAN BATAS DOMINAN

Kalimat	Kelas	Probabilitas	Persentase
kurang ajar siapa yang berani melakukan ini, apa jangan-jangan itu hantu ya	Senang	0,01187531	1,19%
	Sedih	0,03582400	3,58%
	Marah	0,54428500	54,43%
	Takut	0,22563100	22,56%
	Jijik	0,01217200	1,22%
	Terkejut	0,17021300	17,02%

TABEL VIII
EKSPERIMEN NILAI BATAS DOMINAN

$Min(p)$	β	d	Emosi Dominan
1,19%	0%	1,19%	Sedih, Marah, Takut, Jijik, Terkejut
1,19%	1%	2,19%	Sedih, Marah, Takut, Terkejut
1,19%	2%	3,19%	Sedih, Marah, Takut, Terkejut
1,19%	3%	4,19%	Marah, Takut, Terkejut
1,19%	4%	5,19%	Marah, Takut, Terkejut
1,19%	5%	6,19%	Marah, Takut, Terkejut
1,19%	6%	7,19%	Marah, Takut, Terkejut
1,19%	7%	8,19%	Marah, Takut, Terkejut
1,19%	8%	9,19%	Marah, Takut, Terkejut
1,19%	9%	10,19%	Marah, Takut, Terkejut
1,19%	10%	11,19%	Marah, Takut, Terkejut

mengenali nilai dominan pada setiap kelas emosi. Penentuan kelas emosi dominan menggunakan persamaan batas dominan dari nilai persentase pada Tabel VI.

C. Penentuan Emosi Dominan

Persamaan batas dominan yang ditunjukkan pada (3) belum mendapatkan nilai β yang optimal. Untuk mendapatkan nilai β yang optimal, digunakan salah satu kalimat uji pada Tabel IV untuk dilakukan pengujian. Kalimat tersebut ditunjukkan pada Tabel VII dengan nilai probabilitas dan persentase yang dihasilkan model *multinomial naïve Bayes*.

Dari nilai persentase yang ditunjukkan pada Tabel VII kemudian diambil kelas dengan nilai terkecil sebagai nilai parameter $min(p)$ pada persamaan batas dominan. Selanjutnya, dilakukan eksperimen menggunakan nilai $\beta = 0-10\%$ untuk menentukan nilai d sebagai nilai batas dominan dan emosi dominan yang dihasilkan. Dalam hal ini, penentuan emosi dominan ditentukan dengan syarat nilai persentase emosi \geq nilai batas dominan. Hasil eksperimen tersebut ditunjukkan pada Tabel VIII.

Terlihat pada Tabel VIII bahwa nilai $min(p)$ yang didapat dari Tabel VII adalah 1,19%. Kemudian, nilai tersebut ditambahkan dengan nilai β untuk menghasilkan nilai d sebagai nilai batas dominan. Hasil eksperimen pada Tabel VIII menunjukkan bahwa nilai $\beta = 0\%$ menghasilkan lima emosi dominan dan nilai $\beta = 1-2\%$ menghasilkan empat emosi dominan. Kemudian, nilai $\beta = 3\%$ menghasilkan tiga emosi dominan dan tidak berubah jumlah emosi dominannya hingga nilai $\beta = 10\%$. Hal ini menunjukkan bahwa pada metode persamaan batas dominan, dapat digunakan nilai $\beta = 3-10\%$,

TABEL IX
EMOSI DOMINAN KALIMAT UJI

No.	Batas Dominan		Emosi Dominan
	Probabilitas	Persentase	
1.	0,041934	4,19%	Marah, Takut, Terkejut
2.	0,037688	3,76%	Senang, Takut
3.	0,079944	7,99%	Sedih, Takut
4.	0,035331	3,53%	Senang, Sedih, Terkejut
5.	0,047056	4,70%	Sedih, Marah
6.	0,058295	5,83%	Marah, Takut
7.	0,030841	3,09%	Takut, Terkejut
8.	0,058368	5,84%	Sedih, Terkejut
9.	0,056827	5,69%	Senang, Sedih
10.	0,030664	3,07%	Takut
11.	0,194420	19,47%	-

TABEL X
EKSPERIMEN PADA KALIMAT TUNGGAL

Kalimat	Kelas	Probabilitas	Persentase
Malam ini malam jumat kliwon Cintya tidak berani tidur sendirian	Senang	0,00068286	0,07%
	Sedih	0,00068900	0,07%
	Marah	0,00139000	0,14%
	Takut	0,99583000	99,58%
	Jijik	0,00070000	0,07%
	Terkejut	0,00070000	0,07%

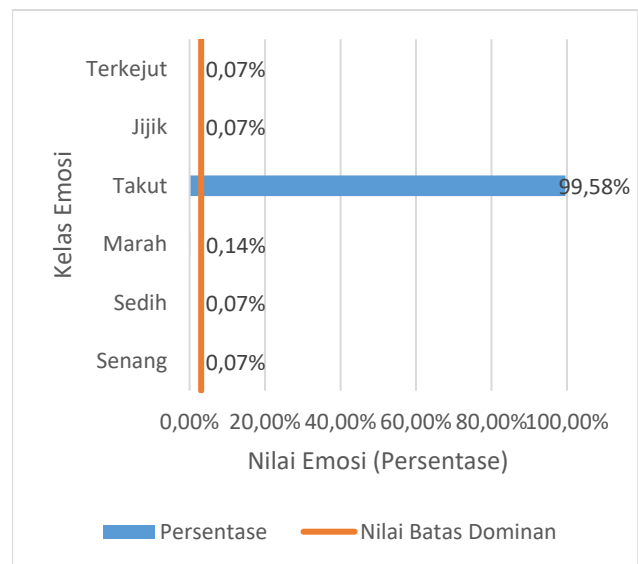
karena emosi dominan yang dihasilkan tidak berubah lagi. Oleh karena itu, pada makalah ini diambil nilai $\beta = 3\%$ sebagai nilai terkecil pada indeks tersebut. Dengan nilai $\beta = 3\%$ pada persamaan batas dominan, dapat dihasilkan emosi dominan dari seluruh kalimat uji pada Tabel IV yang ditunjukkan pada Tabel IX.

Terlihat pada Tabel IX bahwa persamaan batas dominan dapat menghasilkan satu sampai tiga emosi dominan pada kalimat majemuk (kalimat 1-9) dan menghasilkan satu emosi dominan pada kalimat tunggal (kalimat 10). Namun, persamaan batas dominan tidak menghasilkan emosi dominan kalimat bersifat netral (kalimat 11). Hal ini terjadi karena kalimat yang bersifat netral tidak memiliki ekspresi emosi yang dominan atau dapat dikatakan bahwa emosi pada kalimat bersifat netral bergantung pada sudut pandang pembaca kalimat tersebut.

D. Pemetaan Emosi

Pada makalah ini, proses pemetaan emosi dilakukan menggunakan klasifikasi teks dengan model *multinomial naïve Bayes* yang disertai metode persamaan batas dominan. Namun, untuk menunjukkan optimasi dari metode persamaan batas dominan pada pemetaan emosi, berikut ditunjukkan perbandingan hasil pemetaan emosi tanpa persamaan batas dominan dan pemetaan emosi menggunakan persamaan batas dominan pada tiga jenis kalimat (kalimat tunggal, kalimat majemuk, kalimat bersifat netral) yang diambil dari Tabel IV.

Pertama, ditunjukkan hasil eksperimen pada kalimat tunggal yang diambil dari Tabel IV (kalimat 10), dengan nilai probabilitas dan nilai persentase yang ditunjukkan pada Tabel



Gbr. 3 Grafik pemetaan emosi pada kalimat tunggal.

TABEL XI
PERBANDINGAN PEMETAAN EMOSI DOMINAN PADA KALIMAT TUNGGAL

Kalimat	Pemetaan Emosi	
	MNB	MNB + PBD
Malam ini malam jumat kliwon Cintya tidak berani tidur sendirian.	Takut (99,58%)	Takut (99,58%)

MNB = *Multinomial Naïve Bayes*; PBD = *Persamaan Batas Dominan*

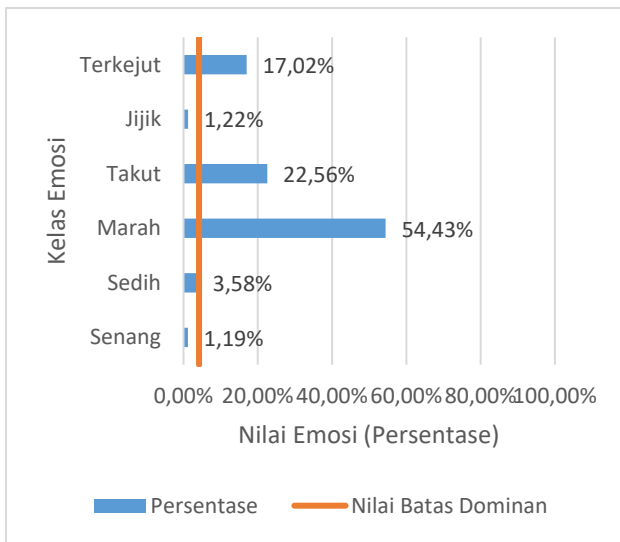
X dan grafik pemetaan emosi yang ditunjukkan pada Gbr. 3. Kemudian, ditunjukkan perbandingan hasil pemetaan emosi tanpa persamaan batas dominan dengan pemetaan emosi menggunakan persamaan batas dominan pada Tabel XI.

Pada Gbr. 3, garis horizontal berwarna biru menunjukkan kelas emosi dan garis vertikal berwarna oranye menunjukkan nilai batas dominan dengan nilai persentase 3,07%. Terlihat hanya kelas emosi takut yang memiliki nilai di atas nilai batas dominan, dengan persentase 99,58%, sedangkan nilai pada kelas lain di bawah nilai batas dominan. Dapat disimpulkan bahwa kalimat tersebut hanya memiliki emosi takut yang dominan tanpa diikuti emosi lain.

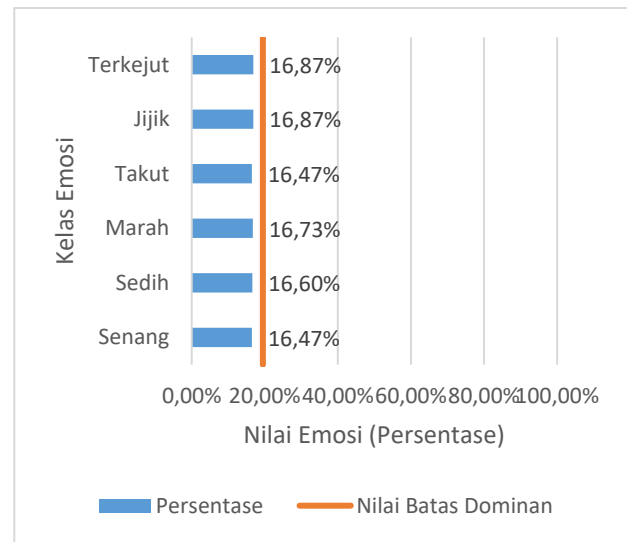
Kemudian, pada Tabel XI tampak bahwa pemetaan emosi tanpa persamaan batas dominan dan pemetaan emosi menggunakan persamaan batas dominan menghasilkan satu emosi yang sama, yaitu emosi takut dengan intensitas emosi 98,58%. Dalam hal ini, persamaan batas dominan tidak menunjukkan hasil yang berbeda pada pemetaan emosi kalimat tunggal.

Kedua, ditunjukkan hasil eksperimen pada kalimat majemuk yang diambil dari Tabel IV (kalimat 1) dengan nilai probabilitas dan nilai persentase ditunjukkan pada Tabel XII dan grafik pemetaan emosi ditunjukkan pada Gbr. 4. Lalu, ditunjukkan perbandingan hasil pemetaan emosi tanpa persamaan batas dominan dengan pemetaan emosi menggunakan persamaan batas dominan pada Tabel XIII.

Pada Gbr. 4, garis horizontal berwarna biru menunjukkan kelas emosi dan garis vertikal berwarna oranye menunjukkan nilai batas dominan dengan nilai persentase 4,19%. Terdapat tiga kelas emosi dengan nilai di atas nilai batas dominan, yaitu



Gbr. 4 Grafik pemetaan emosi pada kalimat majemuk.



Gbr. 5 Grafik pemetaan emosi pada kalimat bersifat netral.

TABEL XII
EKSPERIMEN PADA KALIMAT MAJEMUK

Kalimat	Kelas	Probabilitas	Persentase
kurang ajar siapa yang berani melakukan ini, apa jangan-jangan itu hantu ya	Senang	0,01187531	1,19%
	Sedih	0,03582400	3,58%
	Marah	0,54429000	54,43%
	Takut	0,22563000	22,56%
	Jijik	0,01217000	1,22%
	Terkejut	0,17021000	17,04%

TABEL XIII
PERBANDINGAN PEMETAAN EMOSI DOMINAN PADA KALIMAT MAJEMUK

Kalimat	Pemetaan Emosi	
	MNB	MNB + PBD
kurang ajar siapa yang berani melakukan ini, apa jangan-jangan itu hantu ya	Marah (54,43%)	Marah (54,43%), Takut (22,56%), Terkejut (17,04%)

MNB = *Multinomial Naïve Bayes*; PBD = *Persamaan Batas Dominan*

kelas emosi marah dengan persentase 54,42%, kelas emosi takut dengan persentase 22,55%, dan kelas emosi terkejut dengan persentase 17,04%. Kelas-kelas lainnya memiliki nilai persentase di bawah nilai batas dominan. Dapat disimpulkan bahwa kalimat tersebut memiliki emosi dominan marah, takut, dan terkejut.

Terlihat pada Tabel XIII bahwa pemetaan emosi tanpa persamaan batas dominan menghasilkan kelas emosi “marah” dengan intensitas 54,43%, sedangkan pemetaan emosi menggunakan persamaan batas dominan menghasilkan tiga kelas emosi, yaitu marah 54,43%, takut 22,56%, dan terkejut 17,04%. Hal ini menunjukkan bahwa persamaan batas dominan berhasil memetakan emosi pada kalimat majemuk ke dalam beberapa kelas emosi yang dominan.

Ketiga, ditunjukkan hasil eksperimen dari kalimat bersifat netral yang diambil dari Tabel IV (kalimat 11) dengan nilai probabilitas dan nilai persentase ditunjukkan pada Tabel XIV dan grafik pemetaan emosi ditunjukkan pada Gbr. 5. Selanjutnya, ditunjukkan perbandingan hasil pemetaan emosi tanpa

TABEL XIV
EKSPERIMEN PADA KALIMAT NETRAL

Kalimat	Kelas	Probabilitas	Persentase
Presiden Jokowi hari ini melantik menteri baru bernama Handoko Budianto	Senang	0,16467463	16,47%
	Sedih	0,16600300	16,60%
	Marah	0,16733000	16,73%
	Takut	0,16467000	16,47%
	Jijik	0,16866000	16,87%
	Terkejut	0,16866000	16,87%

TABEL XV
PERBANDINGAN PEMETAAN EMOSI PADA KALIMAT NETRAL

Kalimat	Pemetaan Emosi	
	MNB	MNB + PBD
Presiden Jokowi hari ini melantik menteri baru bernama Handoko Budianto.	Jijik (16,87%) / Terkejut (16,87%)	-

MNB = *Multinomial Naïve Bayes*; PBD = *Persamaan Batas Dominan*

persamaan batas dominan dengan pemetaan emosi menggunakan persamaan batas dominan pada Tabel XV.

Pada Gbr. 5, garis horizontal berwarna biru menunjukkan kelas emosi dan garis vertikal berwarna oranye menunjukkan nilai batas dominan dengan nilai persentase 19,43%. Terlihat bahwa tidak ada kelas dengan nilai di atas nilai batas dominan. Setiap kelas emosi pada kalimat tersebut memiliki nilai antara 16,43%-16,91%. Dengan demikian, dapat disimpulkan bahwa pada kalimat netral tidak terdapat emosi yang dominan. Nilai pada kalimat netral cenderung sama atau rata pada grafik pemetaan.

Kemudian pada Tabel XV terlihat bahwa pemetaan emosi tanpa persamaan batas dominan menghasilkan kelas emosi jijik 16,87% atau terkejut 16,87%. Hal ini terjadi karena pemetaan emosi tanpa batas dominan hanya mengandalkan klasifikasi teks menggunakan model *multinomial naïve Bayes* yang mengambil kelas dengan nilai tertinggi. Pada kalimat tersebut terdapat dua kelas emosi dengan nilai sama, sehingga pemetaan

emosi tanpa persamaan batas dominan mengambil salah satu dari kedua emosi tertinggi tersebut, sedangkan pemetaan emosi menggunakan persamaan batas dominan tidak menghasilkan kelas emosi apapun. Hal ini sama dengan yang dipaparkan pada makalah ini, bahwa kalimat bersifat netral tidak memiliki ekspresi emosi.

V. KESIMPULAN DAN SARAN

Makalah ini bertujuan memetakan jenis dan intensitas emosi yang sesuai dari setiap kalimat bahasa Indonesia. Hasil dari pemetaan emosi digunakan untuk menentukan ekspresi wajah yang sesuai dengan kalimat ucapan pada karakter virtual. Hasil eksperimen menunjukkan bahwa pemetaan emosi berdasarkan klasifikasi teks menggunakan model *multinomial naïve Bayes* yang disertai dengan metode persamaan batas dominan dapat memetakan emosi pada kalimat majemuk ke dalam beberapa kelas emosi. Hal ini dibuktikan dengan hasil eksperimen pada kalimat majemuk yang menunjukkan lebih dari satu emosi dominan dengan nilai intensitas berupa persentase yang berbeda-beda. Kemudian, pemetaan emosi pada kalimat tunggal tetap menghasilkan satu emosi seperti pada proses klasifikasi teks, sedangkan pada pemetaan emosi pada kalimat bersifat netral, metode persamaan batas dominan dapat mengenali kalimat bersifat netral tanpa menambahkan kelas netral. Emosi netral dibuktikan dengan hasil eksperimen yang menunjukkan bahwa intensitas emosi pada kalimat bersifat netral cenderung sama atau rata pada grafik pemetaan. Hal ini terjadi karena kalimat yang bersifat netral tidak memiliki ekspresi emosi yang dominan, atau dapat dikatakan bahwa emosi pada kalimat bersifat netral tergantung pada sudut pandang pembaca kalimat tersebut.

Meskipun pada makalah ini telah dihasilkan pemetaan emosi dominan yang sesuai dari kalimat bahasa Indonesia, tetapi pemetaan emosi dominan masih berupa jenis dan intensitas emosi saja. Untuk dapat menghasilkan ekspresi wajah berdasarkan pemetaan emosi dominan, perlu dilakukan penelitian lebih lanjut tentang *Facial Action Coding System* (FACS). FACS adalah sistem berbasis anatomi yang dikembangkan oleh Paul Ekman untuk menggambarkan gerakan wajah berdasarkan *action unit*. Setiap *action unit* mewakili gerakan otot wajah yang berbeda.

REFERENSI

- [1] *World Population Prospects 2019*, United Nations, 2019.
- [2] (2019) "Perkembangan Pariwisata Oktober 2019" [Online], <https://www.bps.go.id/pressrelease/2019/12/02/1618/jumlah-kunjungan-wisman-ke-indonesia-oktober-2019-mencapai-1-35-juta-kunjungan-.html>, tanggal akses: 01-Mar-2020.
- [3] (2019) "Tenaga Kerja Asing di RI Meroket 38%, Terbanyak dari China!" [Online], <https://www.cnbcindonesia.com/news/20190908075511-4-97843/tenaga-kerja-asing-di-ri-meroket-38-terbanyak-dari-china>, tanggal akses: 01-Mar-2020.
- [4] S. Widodo (2013) "Bahasa Indonesia Menuju Bahasa Internasional" [Online], <http://badanbahasa.kemdikbud.go.id/lamanbahasa/content/bahasa-indonesia-menuju-bahasa-internasional>, tanggal akses: 01-Mar-2020.
- [5] Aripin, H. Haryanto, dan S. Sumpeno, "A Realistic Visual Speech Synthesis for Indonesian Using a Combination of Morphing Viseme and Syllable Concatenation Approach to Support Pronunciation Learning," *Int. J. Emerg. Technol. Learn.*, Vol. 13, No. 8, hal. 19–37, 2018.
- [6] Arifin, Muljono, S. Sumpeno, dan M. Hariadi, "Towards Building Indonesian Viseme: A Clustering-Based Approach," *Proceeding - IEEE CYBERNETICSCOM 2013: IEEE International Conference on Computational Intelligence and Cybernetics*, 2013, hal. 57–61.
- [7] Arifin, S. Sumpeno, M. Hariadi, dan H. Haryanto, "A Text-to-Audiovisual Synthesizer for Indonesian by Morphing Viseme," *Int. Rev. Comput. Softw.*, Vol. 10, No. 11, hal. 1149–1156, Nov. 2015.
- [8] Arifin, S. Sumpeno, Muljono, dan M. Hariadi, "A Model of Indonesian Dynamic Visemes from Facial Motion Capture Database Using a Clustering-Based Approach," *IAENG Int. J. Comput. Sci.*, Vol. 44, No. 1, hal. 41–51, 2017.
- [9] H. Haryanto dan Aripin, "A Finite State Machine Model to Determine Syllables of Indonesian Text," *2019 1st International Conference on Cybernetics and Intelligent System (ICORIS)*, 2019, hal. 238–241.
- [10] S.S.T.W. Sasangka, *Seri Penyuluhan Bahasa Indonesia: Kalimat*, Jakarta, Indonesia: Pusat Pembinaan dan Pemasarakatan Badan Pengembangan dan Pembinaan Bahasa Kementerian Pendidikan dan Kebudayaan, 2014.
- [11] N.H. Frijda, "Moods, Emotion Episodes, and Emotions," dalam *Handbook of Emotions*, M. Lewis dan J.M. Haviland, Eds., New York, USA: Guilford Press, 1993, hal. 381–403.
- [12] P. Ekman dan D. Cordaro, "What is Meant by Calling Emotions Basic," *Emot. Rev.*, Vol. 3, No. 4, hal. 364–370, Okt. 2011.
- [13] M. Arief, M.I. Wardhana, S. Sumpeno, dan M. Hariadi, "Emotion Expression of Three Dimensional Face Model Using Naïve Bayes and Fuzzy Logic," *Int. J. Comput. Sci. Netw. Secur.*, Vol. 10, No. 5, hal. 208–214, 2010.
- [14] J.E. Prawitasllri, "Mengenali Emosi Melalui Komunikasi Nonverbal," *Bul. Psikol.*, Vol. 3, No. 1, hal. 27–43, 2016.
- [15] E. Martiana, R.Y. Hakkun, N. Rosyid M., dan M. Firodh, "Machine Learning dalam Program Chatting untuk Merespon Emosi Teks Berbahasa Indonesia Menggunakan Text Mining dan Naïve Bayes," *The 13th Industrial Electronics Seminar 2011 (IES 2011)*, 2011, hal. 325–330.
- [16] Arifin, "Classification of Emotions in Indonesian Text using K-NN Method," *Int. J. Inf. Electron. Eng.*, Vol. 2, No. 6, hal. 899–903, 2012.
- [17] R. Sitorus dan H.S. Dachlan, "Analisis Pengaruh Frasa pada Deteksi Emosi Dari Teks Menggunakan Vector Space Model," *J. EECCIS*, Vol. 11, No. 1, hal. 41–47, 2017.
- [18] M.C. Noviardini, A.B. Osmond, dan C. Setianingsih, "Klasifikasi Emosi pada Lirik Lagu menggunakan Metode Naïve Bayes Classifier," *e-Proceeding of Engineering*, Vol. 5, No. 3, hal. 6187–6194, 2018.
- [19] A. Librian (2015) "High Quality Stemmer Library for Indonesian Language (Bahasa)" [Online], <https://github.com/sastrawi/sastrawi>, tanggal akses: 23-Apr-2020.
- [20] A. Juan dan H. Ney, "Reversing and Smoothing the Multinomial Naïve Bayes Text Classifier," *Proceedings of the 2nd Int. Workshop on Pattern Recognition in Information Systems (PRIS 2002)*, 2002, hal. 200–212.
- [21] J. Bai dan J.-Y. Nie, "Using Language Models for Text Classification," *Proc. of the Asia Information Retrieval Symposium*, 2004, hal. 1–6.