



PENDETEKSI PENGENALAN EMOSI PADA MANUSIA MENGUNAKAN HIDDEN MARKOV MODEL DAN BIDIRECTIONAL ASSOCIATIVE MEMORY DENGAN SUARA

Rini Meiyanti¹⁾, Cut Lika Mestika Sandy²⁾

^{1 2)}Jurusan Teknik Informatika, Fakultas Komputer dan Multimedia,
Universitas Islam Kebangsaan Indonesia
e-mail: [rinimeiyanti522@gmail.com](mailto:rinameiyanti522@gmail.com)

Abstract

[Detection of Human Emotion Recognition using Hidden Markov Model and Bidirectional Associative Memory with voice] A person's emotional control system through voice can use the Hidden Markov Model (HMM) algorithm. However, to see if the performance of the HMM algorithm in the application system is optimal or not, a comparison is needed in order to obtain maximum results. Therefore, the researchers performed the performance of recognizing a person's emotions using the HMM algorithm and the Bidirectional Associative Memory (BAM) algorithm through voice. The Hidden Markov Model (HMM) consists of a Markov chain in the first part that hides the state, therefore the internal behavior of the model remains invisible. While the BAM algorithm can process incomplete input, because of the reciprocal relationship between the output layer to the input layer. In the BAM algorithm, the value of the test sound and the value of the training voice sample obtained will be searched for the vector value using a weight value search which is done by changing the binary matrix into a bipolar matrix. In this study, we will create an application system that can detect sounds in the form of angry, happy, and neutral emotions. And the database used is the sound of film recordings. This research was conducted to produce a system that can recognize the probability of emotions in the angry, happy and neutral categories, namely by showing the performance of the two methods so that we can find out which method produces the maximum output.

Keywords: HMM; BAM; Emotions; Voice; Detection.

Abstrak

Sistem pengendalian emosi seseorang melalui suara dapat menggunakan algoritma *Hidden Markov Model* (HMM). Namun, untuk melihat kinerja algoritma HMM dalam sistem aplikasi sudah optimal atau belum, diperlukan suatu perbandingan agar memperoleh hasil yang lebih maksimal. Maka dari itu peneliti melakukan unjuk kerja pengenalan emosi seseorang dengan menggunakan algoritma HMM dan algoritma *Bidirectional Associative Memory* (BAM) melalui suara. *Hidden Markov Model* (HMM) terdiri dari rantai markov pada bagian pertama yang menyembunyikan state, oleh karena itu perilaku internal model tetap tidak terlihat. Sedangkan algoritma BAM dapat memproses input yang tidak lengkap, karena adanya hubungan timbal balik antara dari lapisan output ke lapisan input. Pada algoritma BAM, nilai suara pengujian dan nilai sampel suara pelatihan yang diperoleh akan dicari nilai vektornya menggunakan pencarian nilai bobot yang dilakukan dengan cara mengubah matriks biner ke dalam matriks bipolar. Pada penelitian ini akan membuat sebuah sistem aplikasi yang dapat mendeteksi suara dalam bentuk emosi marah, bahagia, dan netral. Dan database yang digunakan adalah suara dari rekaman film. Penelitian ini dilakukan untuk menghasilkan sistem yang dapat mengenali probabilitas emosi pada kategori marah, bahagia dan netral, yaitu dengan menunjukkan unjuk kerja dari kedua metode sehingga kita dapat mengetahui metode mana menghasilkan output yang maksimal.

Kata Kunci: HMM; BAM; Emosi; Suara; Deteksi.

1. Pendahuluan

Emosi mewarnai bahasa dan bertindak sesuai kebutuhan bahan sebagai dua cara alami dari manusia ke manusia komunikasi dan interaksi. Sebagai pendengar kita juga bereaksi dengan keadaan emosi pembicara dan menyesuaikan perilaku kita tergantung pada jenis emosi yang ditransmisikan oleh pembicara. Kemajuan teknologi terkini telah memungkinkan manusia untuk berinteraksi dengan komputer melalui modalitas non-tradisional (misalnya. *keyboard*, *mouse*) seperti suara, gerakan, ekspresi wajah dan lain sebagainya. Interaksi ini masih kekurangan dalam komponen emosi. Dikatakan bahwa untuk benar-benar mencapai komputer manusia afektif interaksi cerdas ada kebutuhan untuk komputer supaya dapat berinteraksi secara alami dengan pengguna, serupa dengan cara interaksi manusia berlangsung seperti biasanya. Beberapa penelitian telah dilakukan yang terdiri dari manusia klasik interaksi dan interaksi manusia dan komputer. Mereka menyimpulkan bahwa untuk interaksi emosi yang cerdas memainkan sebuah bahan penting. Seorang bayi belajar mengenali emosi informasi sebelum memahami informasi semantik pada ucapan ibunya.

Suara merupakan suatu media yang bisa digunakan untuk mengekspresikan suatu emosi. Kita dapat mengidentifikasi banyak jenis emosi dengan berbicara atau mendengarkan suara. Hal ini disebut juga dengan sinyal suara. Dengan melihat orang berbicara dengan intonasi yang berbeda, maka akan berbeda juga dalam mengekspresikan emosi. Dengan melihat juga mendengar seseorang berbicara kita dapat mengetahui bagaimana emosi seseorang tersebut dengan intonasi saat mereka berbicara.

HMM adalah pendekatan stokastik yang memodelkan masalah yang diberikan sebagai “proses stokastik ganda” dimana data yang diamati dianggap sebagai hasil dari melewati proses “benar” (tersembunyi) melalui proses kedua. Kedua proses tersebut harus dicirikan hanya dengan menggunakan salah satu yang dapat diamati (Nilsson & Ejnarsson, 2002).

BAM terdiri dari *neuron* yang tersusun dalam dua lapisan X dan Y. *Neuron-neuron* dalam satu lapisan sepenuhnya saling berhubungan dengan neuron pada lapisan kedua. Tidak ada interkoneksi antar neuron pada lapisan yang sama. Bobot dari lapisan X ke lapisan Y sama dengan bobot dari lapisan Y ke lapisan X. Dinamika melibatkan dua lapisan interaksi. Karena memori memproses informasi dalam waktu dan melibatkan aliran data dua arah, pada prinsipnya berbeda dari asosiasi linier, meskipun kedua jaringandigunakan untuk menyimpan pasangan asosiasi. Ini juga berbeda dari memori auto-asosiatif berulang dalam pembaruannya modus (Jabr & Kareem, 2015).

Berdasarkan permasalahan tersebut, maka perlu di lakukan unjuk kerja sistem aplikasi pengenalan jenis emosi dengan menggunakan algoritma *Hidden Markov Model* (HMM) dan *Bidirectional Associative Memory* (BAM) pada suara. Dengan menggunakan beberapa sampel suara rekaman potongan-potongan dari film sehingga menghasilkan pengenalan suara dalam bentuk emosi marah, bahagia dan netral, maka suatu sistem dapat membedakan jenis suara yang dikenali dengan berbagai resister suara sehingga akan diketahui probabilitas emosi juga seberapa akurat kedua algoritma tersebut dalam mendeteksi jenis emosi pada suara.

2. Metode

A. Emosi

Proses mengenali emosi bicara untuk klasifikasi emosi dalam penelitian ini mengikuti beberapa tahapan. Proses utama yang paling penting adalah ekstraksi fitur dan klasifikasi. Hasil dari ekstraksi ciri sangat berpengaruh dalam menentukan hasil pengenalan emosi bicara di tahap klasifikasi. Manusia menyampaikan perasaan multidimensi. Salah satunya dalam menentukan emosi seseorang yaitu dengan mendeteksi suara saat seseorang tersebut sedang berbicara. Emosi semacam itu dapat diklasifikasikan dengan berbagai model emosi. Model emosi yang efektif perlu dipilih untuk mendeteksi dan menginterpretasikan emosi dari informasi. Ini akan menggambarkan kumpulan emosi yang relevan dengan pertanyaan tertentu.

B. *Hidden Markov Model* (HMM)

HMM adalah model Markov statistik di mana sistem yang dimodelkan diasumsikan menjadi proses *Markov* dengan status tak teramati (tersembunyi). HMM dapat dianggap sebagai jaringan Bayesian dinamis paling sederhana. Dalam model *Markov* yang lebih sederhana (seperti *Markov chain*), status terlihat langsung oleh pengamat. Oleh karena itu, probabilitas transisi status adalah satu-satunya parameter. Proses Markov merupakan proses stokastik dimana dalam sebuah model Markov kita memiliki status observasi fisik sedangkan HMM mengacu pada proses stokastik ganda di mana kita memiliki pengamatan fisik yang terkait dengan keadaan tersembunyi.

Untuk kasus khusus dengan spesifikasi waktu *diskrit*, *first order*, dan *Markov chain*, dapat rumus ketergantungan probabilitas suatu status terhadap status-status sebelumnya disederhanakan hanya tergantung pada tepat satu status sebelumnya, yaitu:

$$P[qt=j|qt-1=i, qt-2=k, \dots] = P[qt=j|qt-1=i]$$

Penyederhanaan lebih lanjut dengan hanya memperhitungkan proses di bagian kanan, bebas terhadap waktu, maka pendekatan aturan probabilitas transisi status:

$$a_{ij} = P[qt=j|qt-1=i]$$

dengan sifat

$$a_{ij} \geq 0, \text{ untuk setiap } j, i$$

$$\sum a_{ij} = 1, \text{ untuk setiap } i, i=1 \text{ sampai } N$$

Karena mematuhi batasan *stochastic standar*.

C. Bidirectional Associative Memory (BAM)

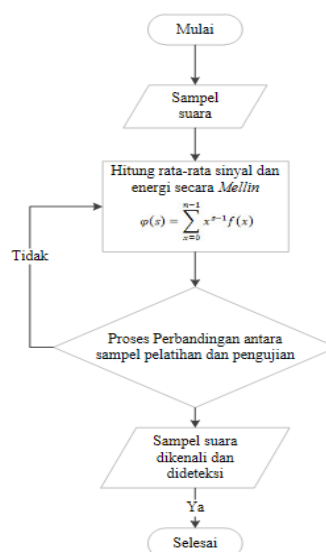
Setiap sistem cerdas harus dapat menerima data dan dapat memproses data tersebut untuk disimpan atau digunakan. Sayangnya tidak semua data jelas dan ringkas, oleh karena itu persyaratan utama untuk setiap sistem cerdas adalah mampu menangani data yang bising atau terdegradasi (Voss, 2007). Mampu mengenali dan mengingat pola bising atau terdegradasi adalah beberapa hal yang dapat dilakukan manusia dengan cepat tetapi masih sulit untuk komputer dan model kecerdasan buatan. Saat ini, jaringan saraf tiruan sedang digunakan untuk masalah ingatan yang bising ini karena kemampuannya untuk mengembangkan penarik untuk setiap pola dan menjadi penyebab stabilitas dan kemampuan beradaptasinya terhadap kebisingan dan degradasi pola. Lebih tepatnya, ingatan asosiatif dua arah digunakan untuk mengasosiasikan dua set pola. Selama bertahun-tahun, beberapa varian model BAM telah diusulkan untuk mengatasi keterbatasan kapasitas penyimpanan model asli dan meningkatkan sensitivitas kebisingannya (Shen & Cruz, 2005). Fungsi transmisi didasarkan pada persamaan Verhulst klasik yang diperluas ke bentuk kubik dengan batas jenuh pada +1 (Chartier et al., 2008). Fungsi transmisi didefinisikan oleh dua persamaan berikut:

$$\forall i, \dots, N, y_i(t+1) = f(\mathbf{W}\mathbf{x}_i(t)) = \begin{cases} 1 & \text{if } \mathbf{W}\mathbf{x}_i(t) > 1 \\ -1 & \text{if } \mathbf{W}\mathbf{x}_i(t) < -1 \\ (\delta + 1)\mathbf{W}\mathbf{x}_i(t) - \delta\mathbf{W}\mathbf{x}_i^3(t) & \text{else} \end{cases} \quad (1)$$

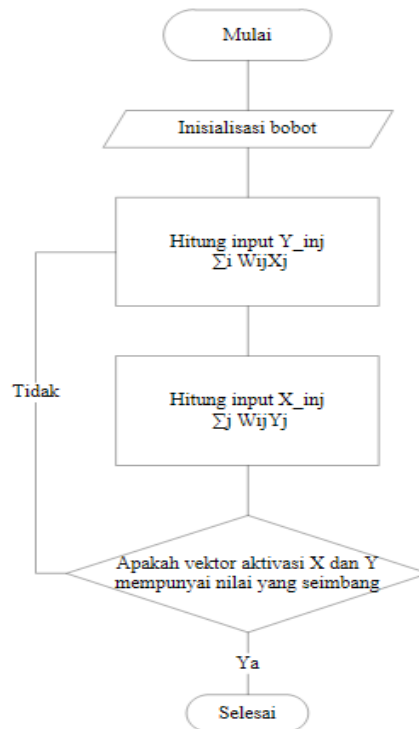
$$\forall i, \dots, M, x_i(t+1) = f(\mathbf{V}\mathbf{y}_i(t)) = \begin{cases} 1 & \text{if } \mathbf{V}\mathbf{y}_i(t) > 1 \\ -1 & \text{if } \mathbf{V}\mathbf{y}_i(t) < -1 \\ (\delta + 1)\mathbf{V}\mathbf{y}_i(t) - \delta\mathbf{V}\mathbf{y}_i^3(t) & \text{else} \end{cases} \quad (2)$$

Dimana N dan M adalah jumlah unit di setiap lapisan. Parameter i adalah indeks dari masing-masing elemen selama pelatihan atau recall. Pada waktu iterasi 1, isi layer di wakikan oleh $x(t)$ dan $y(t)$. *Matriks* bobot adalah W dan V dan δ adalah parameter transmisi umum. Parameter transmisi umum perlu ditetapkan pada nilai antara 0 dan 0,5 untuk memastikan perilaku titik tetap (Chartier et al., 2008). Fungsi transmisi ini digunakan karena tidak memiliki perilaku asimtotik ketika δ berada di antara 0 dan 0,5 dan oleh karena itu dapat digunakan selama pembelajaran dan mengingat. Batas jenuh pada dua penarik, -1 dan 1 memungkinkannya untuk dibandingkan dengan fungsi tipe sigmoid.

D. Skema Sistem Secara Keseluruhan



Gambar 1. Diagram Metode Algoritma *Hidden Markov Model* (HMM)

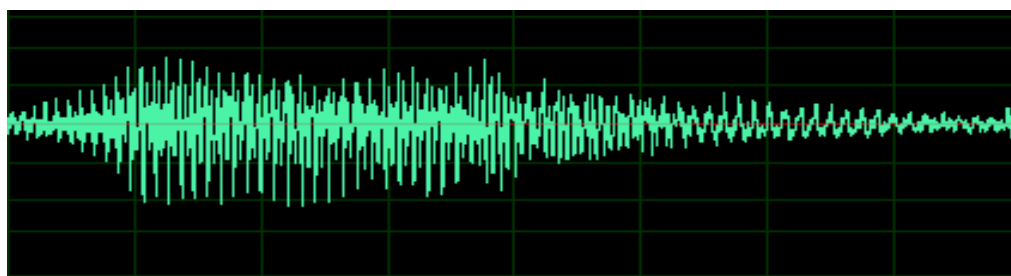


Gambar 2. Diagram Metode Algoritma *Bidirectional Associative Memory* (BAM)

Berdasarkan gambar diatas menunjukkan masing-masing skema sistem dengan menggunakan metode Hidden Markov Model (HMM) dan Bidirectional Associative Memory (BAM). Dalam skema tersebut menjelaskan bahwa setiap sinyal suara akan menerima inputan struktur sampel suara. Struktur sampel suara inputan yang dimaksud adalah sebarisan nilai sinyal suara tepi-tepi objek yang akan mendeteksi jenis suara, dan selanjutnya jaringan melakukan tahapan pelatihan yaitu pengenalan sinyal suara dan tandai ciri, dan uji pengenalan suara.

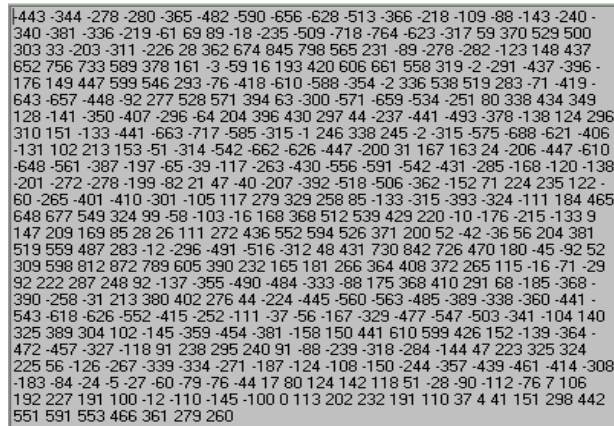
3. Hasil dan Pembahasan

Sampel training dan testing berupa sampel suara secara real-time dan record. Pada penelitian ini akan membuat sebuah sistem aplikasi yang dapat mendeteksi suara dalam bentuk emosi marah, bahagia, dan netral. Untuk sampel suara yang akan diuji akan dilakukan samel suara untuk dilatih terlebih dahulu. Agar nilai sinyal suaranya sudah terecord di dalam sampel suara dilatih. Dan data training yang digunakan adalah suara dari rekaman film. Penelitian ini dilakukan untuk menghasilkan sistem yang dapat mengenali probabilitas emosi pada kategori marah, bahagia dan netral, yaitu dengan menunjukkan unjuk kerja dari kedua metode sehingga kita dapat mengetahui metode mana menghasilkan output yang maksimal.



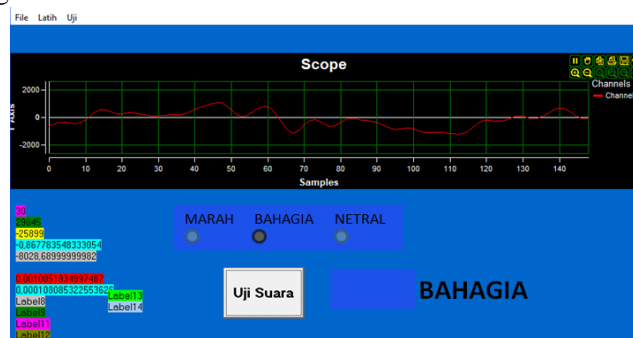
Gambar 3. Mengilustrasikan gambar sinyal suara

Gambar tersebut mengilustrasikan bagaimana gambar pada sinyal suara saat mendeteksi samping suara yang pada saat dilatih dan diuji. Dalam menentukan normalisasi pada sinyal suara yaitu jika nilai $f > 0$ maka $f = 1$, jika $f < 0$ maka $f = -1$ dan jika $f = 0$ maka $f = 0$.

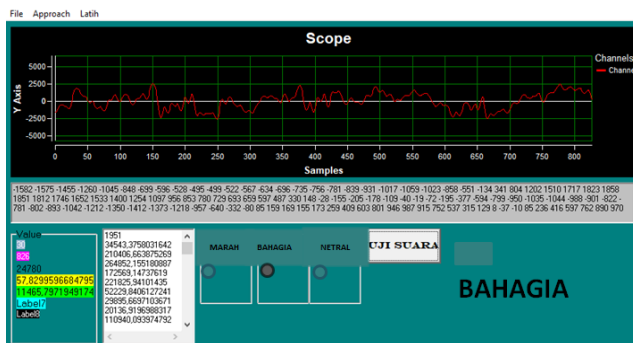


Gambar 4. Nilai sampling suara emosi bahagia

Pelatihan dilakukan dengan menginputkan sampel dari berbagai macam jenis suara emosi berdasarkan register suara emosi yang berjenis marah, bahagia dan netral dengan jumlah sampel pengujian yang memiliki emosi yang berbeda. Hal ini bertujuan untuk melihat sejauh mana sistem dapat mengenali pendeteksian emosi tersebut. Seperti gambar dibawah ini menunjukkan proses pengujian sampel suara dengan menggunakan algoritma HMM dan BAM:



Gambar 5. Tampilan Form Pengujian pada suara jenis Emosional Bahagia menggunakan metode HMM



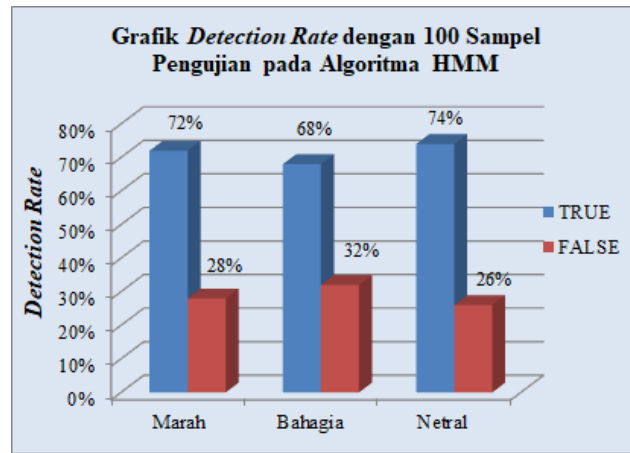
Gambar 6. Tampilan Form Pengujian pada suara jenis Emosional Bahagia menggunakan metode BAM

Pada gambar diatas menampilkan bahwa user melakukan penginputan data sampel suara dengan register suara emosional bahagia dengan menggunakan metode Algoritma HMM dan Algoritma BAM. Setelah dilatih terlebih dahulu untuk setiap sampel suara pada masing-masing emosional yaitu seperti potongan suara marah, bahagia dan netral. Kemudian sistem akan melakukan pengujian setelah dimasukkan potongan suara secara random dan akan mendapatkan hasil detection rate sinyal suara pada gambar diatas menunjukkan bahwa sampel suara tersebut mengandung emosi bahagia.

Tabel berikut menggabungkan empat jenis register suara sekaligus dalam satu tabel, dengan pelatihan sebanyak satu kali untuk masing-masing jenis register dan menggunakan dua sampel pelatihan yang berbeda dengan sampel suara pengujian dan sampel suara *real time* dengan melihat unjuk kerja dari algoritma HMM dan algoritma BAM. Untuk lebih detailnya dapat kita lihat pada tabel berikut :

Tabel 1. Unjuk Kerja Sistem Untuk Proses Pengenalan Suara *Record* pada BAM

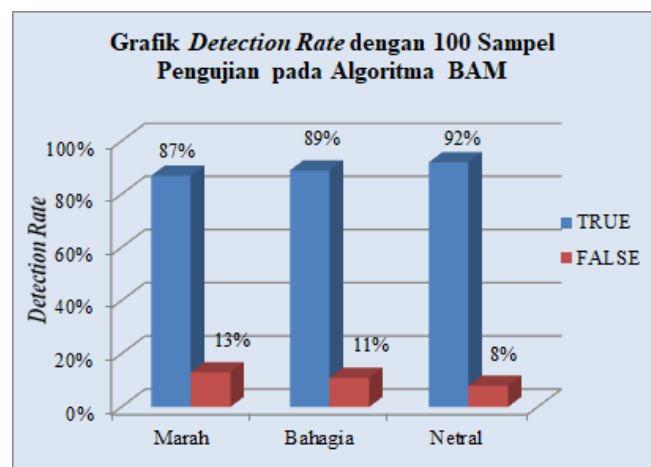
No.	Jumlah Suara Pelatihan	Jumlah Suara Pengujian	Sampel Suara Emosi	Jumlah Pendeteksi yang benar	Detection Rate	
1	50	100	Marah	True Detection	72	0.72
				False Detection	28	0.28
2	50	100	Bahagia	True Detection	68	0.68
				False Detection	32	0.32
3	50	100	Netral	True Detection	74	0.74
				False Detection	26	0.26



Gambar 7. Grafik Detection Rate pada Algoritma HMM

Tabel 1. Unjuk Kerja Sistem Untuk Proses Pengenalan Suara *Record* pada BAM

No.	Jumlah Suara Pelatihan	Jumlah Suara Pengujian	Sampel Suara Emosi	Jumlah Pendeteksi yang benar	Detection Rate	
1	50	100	Marah	True Detection	87	0.87
				False Detection	13	0.13
2	50	100	Bahagia	True Detection	89	0.89
				False Detection	11	0.11
3	50	100	Netral	True Detection	92	0.92
				False Detection	8	0.08



Gambar 8. Grafik Detection Rate pada Algoritma BAM

Pada tabel dan gambar diatas menunjukkan bahwa dari hasil eksekusi program akan didapatkan persentase keberhasilan pengenalan untuk jenis register suara dari beberapa sampel yang di input. Kolom ‘Sampel Suara Pelatihan’ memuat beberapa sampel suara yang berbeda yang dijadikan sebagai pelatihan. Kolom ‘Jumlah Pengujian’ pun demikian, berisi jumlah pengujian yang di lakukan *user*, kemudian ada kolom ‘Identifikasi Benar’, yang berisi data tentang berapa jumlah pengenalan suara yang berhasil disertai dengan persentasenya. Terakhir yaitu kolom ‘Identifikasi Salah’, kebalikan dari kolom ‘Identifikasi Benar’, yang berisi data tentang jumlah pengenalan suara yang tidak berhasil beserta persentasenya.

4. Kesimpulan

Dari hasil penelitian yang dilakukan terhadap beberapa sampel suara, hasil pengenalan suara pada penelitian ini menunjukkan bahwa sistem pendeteksian pengenalan emosi menggunakan algoritma *Hidden Markov Model* (HMM) memiliki kisaran *true detection* sebesar 72% pada jenis emosi marah, 68% pada jenis emosi bahagia dan 74% pada jenis emosi netral. Sedangkan dengan menggunakan algoritma *Bidirectional Associative Memory* (BAM) memiliki kisaran *true detection* sebesar 87 % pada jenis emosi marah, 89% pada jenis emosi bahagia dan 92% pada jenis emosi netral. Persentase *true detection* menunjukkan bahwa algoritma *Hidden Markov Model* (HMM) dan algoritma *Bidirectional Associative Memory* (BAM) berhasil digunakan sebagai metode dalam mendeteksi emosi melalui suara seseorang. Dalam hasil penelitian yang telah dilakukan, pendeteksian yang dilakukan dengan algoritma BAM lebih optimal hasilnya dibandingkan dengan Algoritma HMM. Namun kedua metode tersebut bias digunakan dalam sistem pendeteksi emosi seseorang. Dalam penelitian ini, sistem akan mendeteksi sampel suara yang dilatih dan diuji, dengan begitu maka diharapkan sampel suara yang dimasukkan harus jelas dan tidak terputus putus agar memperoleh hasil yang optimal.

Daftar Pustaka

- Chartier, S., Renaud, P., & Boukadoum, M. (2008). A nonlinear dynamic artificial neural network model of memory. *New Ideas in Psychology*, 26(2), 252–277.
- Jabr, N. A. A., & Kareem, E. I. A. (2015). Modify Bidirectional Associative Memory (MBAM). *International Journal of Modern Trends in Engineering and Research (IJMTER)*, 2(08), 136–151.
- Nilsson, M., & Ejnarsson, M. (2002). *Speech recognition using hidden markov model*.
- Shen, D., & Cruz, J. B. (2005). Encoding strategy for maximum noise tolerance bidirectional associative memory. *IEEE Transactions on Neural Networks*, 16(2), 293–300.
- Voss, P. (2007). Essentials of General Intelligence: The Direct Path to Artificial General Intelligence. In B. Goertzel & C. Pennachin (Eds.), *Artificial General Intelligence* (pp. 131–157). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-68677-4_4