



Hak cipta dan penggunaan kembali:

Lisensi ini mengizinkan setiap orang untuk menggubah, memperbaiki, dan membuat ciptaan turunan bukan untuk kepentingan komersial, selama anda mencantumkan nama penulis dan melisensikan ciptaan turunan dengan syarat yang serupa dengan ciptaan asli.

Copyright and reuse:

This license lets you remix, tweak, and build upon work non-commercially, as long as you credit the origin creator and license it on your new creations under the identical terms.

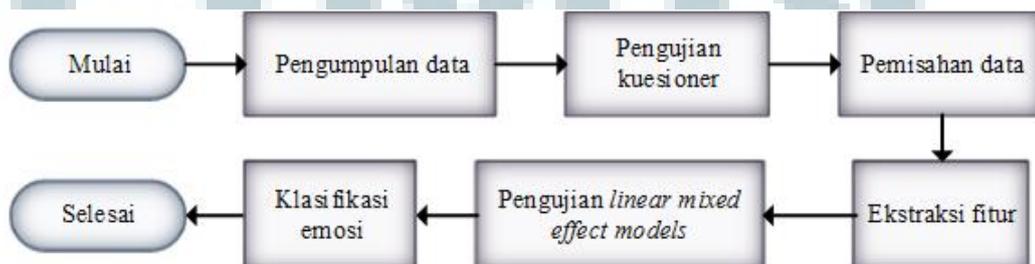
BAB 3

METODE PENELITIAN

3.1 Pengumpulan Data

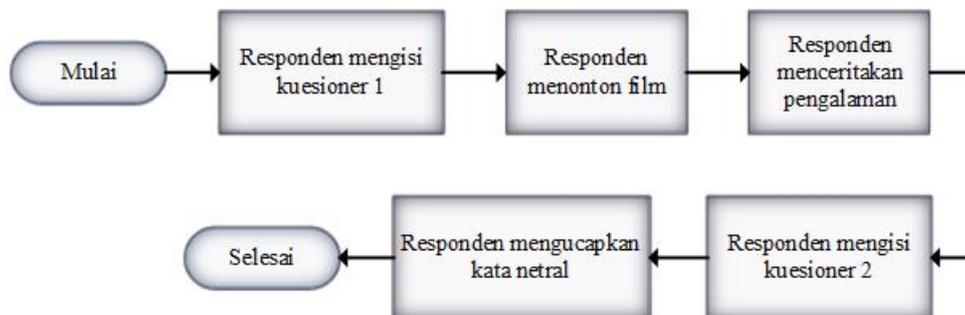
3.1.1 Sumber Data

Data suara untuk emosi senang, sedih, dan netral dikumpulkan dari 40 orang responden, masing-masing 20 orang laki-laki dan perempuan. Beberapa persyaratan untuk menjadi responden adalah berstatus sebagai mahasiswa, berusia 18-24 tahun, dan memiliki cara berbicara yang tidak berlogat. Penulis mengumpulkan responden secara acak dari Universitas Multimedia Nusantara yang memenuhi syarat yang telah disebutkan. Calon responden menyetujui data audio dan video dari penelitian ini dipublikasikan untuk keperluan akademik dengan menandatangani surat persetujuan terlebih dahulu. Responden juga diberi kebebasan untuk mengundurkan diri di tengah proses pengumpulan data sesuai dengan yang tertera di dalam surat persetujuan.



Gambar 3.1 Alur penelitian

3.1.2 Tahap Pengumpulan Data



Gambar 3.2 Alur pengumpulan data setiap emosi

Terdapat beberapa tahap yang dilakukan guna memperoleh data suara yang diperlukan. Sebelum memulai pengumpulan data untuk setiap emosi, responden mengisi kuesioner mengenai kondisi emosinya pada saat itu. Kuesioner ini diadaptasi dari penelitian Ray (2007) untuk memvalidasi film yang digunakan untuk memicu emosi senang dan sedih. Kuesioner berisi pertanyaan apakah responden sudah pernah menonton klip film yang diberikan serta kolom untuk 16 jenis emosi yang terdiri atas emosi (1) terkejut, (2) marah, (3) cemas, (4) tenang, (5) bimbang, (6) tersinggung, (7) jijik, (8) malu, (9) bersemangat, (10) takut, (11) bersalah, (12) senang, (13) tertarik, (14) cinta, (15) bangga, dan (16) sedih. Responden mengisi setiap kolom emosi tersebut dengan menggunakan skala Likert dari 1 apabila emosi tidak dirasakan sama sekali sampai 5 untuk emosi yang sangat dirasakan.

Selanjutnya, responden diminta menonton satu klip film yang disediakan sesuai dengan emosi yang ingin dipicu. Klip film dimanfaatkan sebagai media pemicu emosi karena film memberikan kebebasan pada para peneliti untuk memicu respon

emosi yang kuat, bahkan yang negatif sekalipun, tanpa menyebabkan rasa bahaya atau pelanggaran etis. Berdasarkan penelitian terdahulu, emosi film dinilai nyata dan kuat karena dapat menyebabkan partisipan menangis ketika menyaksikan film sedih, tertawa pada film lucu, dan sebagainya. Meskipun masih terdapat pertanyaan mengenai keabsahan ekologis dari film, film mungkin lebih alami dan umum jika dibandingkan dengan cara lainnya seperti pergerakan wajah yang diarahkan atau hipnosis, yang sangat sedikit menyerupai situasi hidup sehari-hari (Ray, 2007). Terdapat tiga klip film yang disediakan menurut jenis emosi yang paling sesuai, di mana klip tersebut merupakan klip yang digunakan dalam pengembangan *database* film untuk memicu emosi.

Tabel 3.1 Klip film untuk memicu emosi

Emosi	Judul Film	Durasi	Deskripsi Isi
Senang	<i>Wall-E</i>	3 menit 15 detik	Dimulai ketika robot putih terbang maju. Dua robot menari di luar angkasa dan jatuh cinta sambil ditonton oleh orang di dalam pesawat ruang angkasa dan musik dimainkan. Berakhir ketika kedua robot terbang pergi bersama.
Netral	<i>Searching For Bobby Fischer</i>	3 menit 19 detik	Dimulai pada adegan seorang wanita dan anaknya berlari di lorong. Seorang pria menemukan bahwa anak laki-lakinya bisa bermain catur dan bermain dengannya. Berakhir ketika anak laki-laki berdiri dari kursi setelah permainan selesai.

Sedih	<i>Saving Private Ryan</i>	4 menit 14 detik	Adegan dimulai dengan seorang pekerja kantor menulis surat Dear John. Seorang wanita diberitahu bahwa tiga dari empat anaknya meninggal dalam Perang Dunia II. Berakhir dengan wanita tersebut terduduk di berandanya.
-------	----------------------------	---------------------	--

Klip film *Searching for Bobby Fischer* digunakan untuk emosi netral karena memenuhi kriteria yang ditentukan dalam penelitian bersangkutan, di mana rata-rata penilaian untuk semua emosi (kecuali emosi tenang) tidak melebihi 3 dari skala 0-8. Untuk emosi sedih, klip film *Saving Private Ryan* memicu kesedihan dengan intensitas rata-rata 6,7 dan skor 3,2 untuk pencapaian emosi target. Khusus untuk klip emosi senang, *Wall-E*, meskipun tidak dapat memicu emosi senang tanpa gangguan dari emosi non-target, tetapi dapat menimbulkan emosi senang yang diikuti emosi tenang dan cinta (Bartolini, 2011).

Pada tahap berikutnya, responden diminta mengingat pengalaman pribadi yang berhubungan dengan emosi tertentu dan menceritakan pengalaman tersebut kepada penulis. Proses mengingat dan menceritakan pengalaman diharapkan dapat memperkuat emosi yang sebelumnya telah dipicu melalui film yang ditonton (Jurgen, 1996; Parrott & Hertel, 1999). Responden kemudian kembali mengisi kuesioner mengenai kondisi emosi mereka saat itu dan dilanjutkan dengan mengucapkan 30 kata netral yang ditampilkan dalam bentuk presentasi, di mana slide kata telah diatur untuk berpindah setiap 3 detik.

Konten elektronik yang dipublikasikan seperti berita dan majalah elektronik, maupun konten cetak dan yang tidak dipublikasikan telah digunakan untuk membangun *database* SoNaR yang memuat 500 juta kata berbahasa Belanda (Reynaert, Oostdijk, Clercq, Heuvel, & Jong, 2010) sehingga penulis menggunakan 21 artikel dari beberapa situs berita antara lain Detik, Okezone, dan Liputan6 untuk memperoleh daftar kata netral yang bervariasi. Sebanyak 200 kata benda dipilih dari artikel tersebut, kemudian disimpan dalam bentuk dokumen Excel dan disebarakan kepada lima orang yang tidak berpartisipasi sebagai responden. Lima orang tersebut bertugas memberikan penilaian apakah kata-kata yang terdapat dalam daftar termasuk netral atau tidak. Jika kata tersebut dinilai netral oleh minimal tiga orang, maka kata itu dapat menjadi calon kata netral yang digunakan. Pada akhirnya, 30 kata dipilih secara acak dari calon kata yang memenuhi syarat sebagai kata netral.

Tabel 3.2 Daftar kata netral untuk pengumpulan data

No.	Kata Netral
1.	Perut
2.	Mobil
3.	Permen
4.	Hujan
5.	Aktris
6.	Kelapa
7.	Kabin
8.	Garam
9.	Soda
10.	Ekspor
11.	Kamera
12.	Darat
13.	Waktu
14.	Rumah

15.	Ikan
16.	Album
17.	Pasar
18.	Motor
19.	Swasta
20.	Acara
21.	Tangan
22.	Momen
23.	Batik
24.	Negara
25.	Baja
26.	Surat
27.	Calon
28.	Payung
29.	Warga
30.	Daerah

Alat yang diperlukan untuk pengumpulan data adalah mikrofon, laptop, dan kamera video. Mikrofon diletakkan sejauh kurang lebih 25-30 cm dari mulut responden. Perangkat lunak yang digunakan untuk merekam suara adalah Praat dan data suara disimpan dalam format .wav dengan *sampling rate* 16 KHz/16-bit (Joshi & Zalte, 2013). Laptop berperan sebagai sarana untuk menonton klip film, menampilkan daftar kata netral yang perlu diucapkan dan merekam suara. Di samping itu, keseluruhan aktivitas responden selama pengumpulan data direkam oleh kamera video.

3.2 Pengolahan Data

3.2.1 Pengujian Kuesioner

Intensitas setiap emosi pada kuesioner sebelum dan sesudah responden diberikan *treatment* (menonton klip video dan menceritakan pengalaman) untuk masing-masing tahap pengumpulan data emosi senang, sedih, dan netral dibandingkan dengan uji Wilcoxon (lihat Subbab 2.13). Setelah itu, data kuesioner yang diisi oleh responden setelah diberikan *treatment* kembali diuji menggunakan uji Friedman (lihat Subbab 2.12) dan apabila terdapat perbedaan signifikan antara 16 tipe emosi pada kuesioner, dilanjutkan dengan uji *post hoc* menggunakan uji Wilcoxon. Kedua pengujian ini bertujuan untuk mengetahui apakah *treatment* yang diberikan menimbulkan efek terhadap emosi responden dan berhasil memicu emosi target yang diinginkan. Emosi target yang dimaksud adalah emosi senang pada tahap pengumpulan data emosi senang, emosi tenang pada tahap pengumpulan data emosi netral, dan emosi sedih pada tahap pengumpulan data emosi sedih. Uji Friedman dan Wilcoxon dilakukan menggunakan perangkat lunak statistik SPSS.

3.2.2 Pemisahan Data

Data suara yang terkumpul per responden per emosi masing-masing berupa *file* suara berisi pengucapan 30 kata netral secara berturut-turut. Oleh karena itu, data suara dipotong per kata terlebih dahulu menggunakan perangkat lunak Audacity

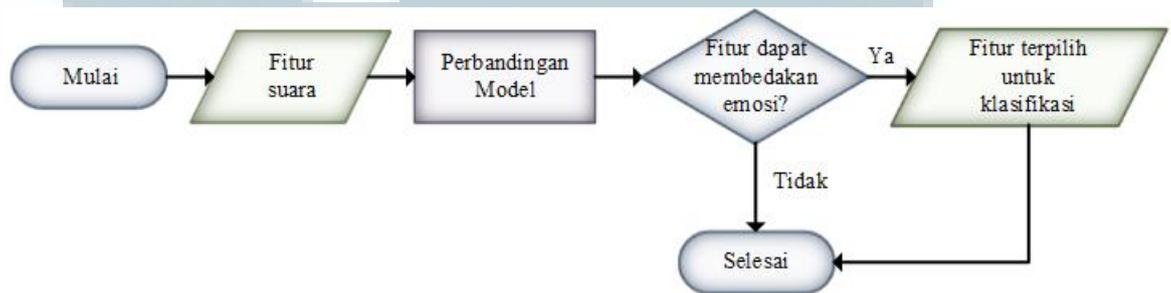
dan disimpan kembali ke dalam format .wav. Selanjutnya, *file* suara dikelompokkan ke dalam *folder* terpisah menurut jenis kelamin responden.

3.2.3 Ekstraksi Fitur

Fitur yang diekstraksi dari data suara responden adalah rata-rata amplitudo, jumlah sampel, dan rata-rata koefisien aproksimasi dari keluarga *wavelet* Daubechies yaitu db1 sampai db4. Amplitudo merupakan pengukuran objektif dari derajat perubahan (positif atau negatif) di dalam tekanan atmosfer yang disebabkan oleh gelombang suara (Hass, 2003) dan menunjukkan volume suara manusia (B'far, 2004). Sinyal suara hasil perekaman merupakan sinyal analog yang telah dikonversi menjadi sinyal digital, di mana $x(n)$ merepresentasikan amplitudo dari sampel ke n dan n merupakan indeks waktu atau angka sampel (Tan & Jiang, 2007) sehingga jumlah sampel menunjukkan durasi dari sinyal suara bersangkutan. *Wavelet* Daubechies db4 sebelumnya telah digunakan untuk mengekstraksi fitur suara dan menghasilkan akurasi pengenalan emosi yang cukup menjanjikan (Joshi & Zalte, 2013) sehingga merupakan keluarga *wavelet* yang digunakan untuk mengekstraksi fitur. Koefisien aproksimasi *wavelet* diperoleh setelah sinyal disaring dengan *low pass* filter yang melewatkan frekuensi rendah dari sinyal dan koefisien dengan energi tinggi (Yadav et al., 2015). Koefisien aproksimasi tingkat satu pada keluarga Daubechies D8, D10, D12, dan D20 memuat lebih dari 96% energi sinyal (Fgee, Phillips, & Robertson, 1999; Yadav et al., 2015), sehingga pada penelitian ini, digunakan dekomposisi tingkat 1 untuk memperoleh koefisien tersebut.

Dari fitur yang ditentukan, ingin diketahui apakah terdapat perbedaan pada amplitudo, jumlah sampel, dan koefisien aproksimasi db1 sampai db4 ketika seseorang berbicara dalam emosi senang, netral dan sedih. Proses ekstraksi fitur suara tiap emosi dilakukan menggunakan bahasa pemrograman Matlab. Fitur yang diekstraksi disimpan dalam dokumen teks berformat .csv dan digunakan pada *linear mixed effect models* untuk mengukur perbedaan yang ditimbulkan oleh emosi dan faktor lainnya pada fitur tersebut.

3.2.4 *Linear Mixed Effect Models*



Gambar 3.3 Alur *linear mixed effect models*

Dengan fungsi *lmer* dalam *library lme4* (Bates, Mächler, Bolker, & Walker, 2015) pada bahasa pemrograman R, *linear mixed effect models* (lihat Subbab 2.6) dibangun untuk menganalisis hubungan antara emosi dengan fitur suara yang diekstraksi. Rata-rata amplitudo, jumlah sampel, dan rata-rata dari koefisien aproksimasi db1 sampai db4 merupakan variabel respon, emosi sebagai *fixed effect*, serta jenis kelamin, subjek, dan kata sebagai *random effect*. Emosi dimasukkan sebagai *fixed effect* dalam model karena fokus utama dalam penelitian ini adalah untuk mengetahui apakah emosi dapat dibedakan melalui fitur suara yang ditentukan. Jenis kela-

min, subjek, dan kata diasumsikan sebagai *random effect* yang dapat menimbulkan variasi pada respon.

Likelihood ratio test (lihat Subbab 2.7) digunakan untuk membandingkan model dengan *fixed effect* emosi dan model tanpa *fixed effect* emosi sehingga dapat diketahui apakah emosi mempengaruhi variabel respon secara signifikan. Perbandingan juga dilakukan dengan model di mana *fixed effect* tetap namun salah satu *random effect* dihilangkan untuk melihat pengaruh yang mungkin diberikan oleh *random effect* tersebut. Fungsi *anova* digunakan untuk melakukan *likelihood ratio test* pada bahasa pemrograman R (Baayen et al., 2008). Apabila emosi menimbulkan perbedaan signifikan pada variabel respon bersangkutan, maka variabel tersebut akan dipilih sebagai fitur masukan untuk proses klasifikasi. Hasil dinyatakan sebagai signifikan apabila nilai p yang diperoleh dari perbandingan menggunakan fungsi *anova* lebih kecil dari 0.05 (*P Values*, n.d.).

3.3 Klasifikasi Emosi

Melalui *linear mixed effect models*, dapat ditentukan fitur suara yang berpotensi untuk digunakan sebagai masukan pada proses klasifikasi. Jika fitur memiliki dimensi yang besar, PCA (lihat Subbab 2.10) diterapkan untuk mengambil 200 elemen pertama dari fitur tersebut. Selanjutnya, klasifikasi emosi dilakukan dengan *neural network* dan/atau *support vector machine (SVM)*. *Neural network* digunakan karena mempunyai performa yang lebih baik dari HMM dan GMM jika jumlah data latih relatif rendah, sedangkan SVM telah banyak digunakan dalam aplikasi

pengenalan pola dan memiliki performa lebih baik daripada *classifier* lain (Ayadi, Kamel, & Karray, 2011).

Struktur dari *neural network* yang dibangun memiliki 1 lapisan masukan, 1 lapisan tersembunyi, dan 1 lapisan keluaran. Jumlah *hidden neuron* yang digunakan untuk *neural network* adalah 1 neuron untuk fitur dengan dimensi 1 serta 30 dan 50 neuron untuk fitur dengan dimensi 200. Jenis *neural network* yang digunakan adalah PatternNet karena memiliki performa yang baik untuk sistem pengenalan (Al-allaf, Tamimi, & AbdAlKader, 2012). SVM dibangun menggunakan kernel RBF, dengan parameter sesuai dengan nilai *default* pada *library* yang digunakan.

Untuk mengecek akurasi dari *classifier*, penulis menerapkan *stratified 10-fold cross validation* karena proses *cross validation* dapat mencegah masalah *overfitting* (Hsu, Chang, & Lin, 2003) dan *10-fold* mengurangi variasi dari akurasi (Kohavi, 1995). Nilai akurasi diperoleh dari merata-ratakan persentase klasifikasi yang benar dari tiap *fold*. Keseluruhan proses klasifikasi dilakukan dengan memanfaatkan fungsi dari *Neural Network Toolbox* dan *Bioinformatics Toolbox* pada bahasa pemrograman Matlab.