



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 III

METODOLOGI PENELITIAN

3.1 Variabel Penelitian

Peneliti menggunakan

- 3.1.1 Variabel independen, yaitu *input* suara.
- 3.1.2 Variabel moderator, yaitu fitur, jenis *learning*, dan kondisi pengujian.
- 3.1.3 Variabel kontrol, yaitu lokasi pengujian.
- 3.1.4 Variabel dependen yaitu *output* suara dari hasil prediksi.

Variabel *input* suara berbentuk rekaman yang didapat dari mikrofon. Variabel fitur berbentuk function yang mengekstrak *input* suara menjadi sekumpulan nilai. Variabel jenis *learning* adalah variabel berbentuk *function* yang mengeksekusi data hasil keluaran fitur dan *database*. Variabel kondisi pengujian, berbentuk kriteria pengujian, seperti jarak suara dari mikrofon di tiap pengujian. Variabel lokasi pengujian mengontrol semua pengujian dilakukan di tempat yang sama. Terakhir, variabel *output* merupakan hasil prediksi dari aplikasi, berbentuk teks yang ditampilkan di bagian akhir dari alur jalannya aplikasi pada tiap *cycle*.

3.2 Populasi dan Sampel

Peneliti menggunakan metode penelitian *factorial design*, dimana peneliti memberi variabel moderator pada variabel independen untuk melakukan observasi pada variabel dependen sebagai hasil. Ketika terdapat perbedaan pengaruh pada hasil prediksi suara, maka penyebab utamanya bukanlah karena perbedaan suara *input*, tetapi karena adanya variabel moderator. Dalam hal ini adalah jenis fitur yang dipasang, jenis *learning* yang dipakai, dan kondisi pengujian yang berbeda-beda. Populasi pada penelitian ini adalah semua objek penelitian yang dilakukan.

Sampel adalah bagian dari populasi. Sedangkan populasi adalah totalitas semua nilai yang mungkin, baik hasil menghitung maupun pengukuran kuantitatif maupun

kualitatif dari karakteristik tertentu mengenai sekumpulan obyek yang lengkap dan jelas.

Bila populasi terlalu besar untuk dilakukan survey secara menyeluruh atau total, maka akan dilakukan sampel, dimana sampel mempersingkat waktu dan biaya dari penelitian itu sendiri, tetapi sampel harus bisa mewakili populasi tersebut. Ada beberapa hal yang bisa ditinjau dan diperhitungkan agar sampel bisa mewakili populasi secara matematis dan representatif:

- 3.2.1 populasi target adalah semua objek/instrumen penelitian,
- 3.2.2 unit sampel/unit dasar yang digunakan adalah macam gelombang suara yang diuji,
- 3.2.3 tingkat kesalahan dan keyakinan sampel yang digunakan adalah 10% dimana tingkat keyakinan mempunyai bobot 90% dari keseluruhan 100%.

3.3 Teknik Sampling

Dari beberapa teknik pengambilan sampel seperti:

3.3.1 Simple Random Sampling

Pada simple random sampling pengambilan sampel dilakukan secara acak dengan metode angka acak tertentu dari seluruh populasi yang ada. Ciri utama sampling ini adalah setiap unsur dari keseluruhan populasi mempunyai kesempatan yang sama untuk dipilih.

3.3.2 Stratified Random Sampling

Pada stratified random sampling pengambilan sampel berdasarkan informasi awal berkaitan dengan stratifikasi dari populasi. Dalam hal ini pengambilan sampel pada setiap stratifikasi dilakukan secara acak, sama halnya seperti yang dilakukan pada simple random sampling. Teknik ini digunakan apabila populasinya heterogen atau terdiri atas kelompok-kelompok yang bertingkat, misalnya menurut pendapatan, usia, jenjang pendidikan.

3.3.3 *Clustered Sampling*

Pada teknik ini total populasi dibagi menjadi sekumpulan cluster unit sample. Selanjutnya masing-masing cluster ditarik sampelnya secara acak. Teknik ini digunakan apabila populasi tersebar di beberapa daerah, provinsi, kabupaten, kecamatan dan seterusnya.

3.3.4 *Systematic Sampling*

Teknik pengambilan sampel pada metode ini dilakukan dengan memilih unit sampel berdasarkan daftar dan penarikannya dilakukan berdasarkan interval tertentu, misalnya setiap kelipatan 5 atau 10 dari daftar pegawai.

Peneliti menggunakan stratified random sampling karena ada pembagian tiap sampel yang berdasarkan jenis bunyi, apakah itu suara pintu, atau suara air. Peneliti mengambil sampel secara acak dari tiap-tiap kelas nantinya.

3.4 Instrumen Penelitian

Skala pengukuran yang digunakan peneliti dalam penelitian adalah skala *ratio* dimana terdapat nilai-nilai rentang pada data yang diolah, dikategorikan, dibuat peringkat, mempunyai jarak, dan bisa dibandingkan karena variabel yang ada berbentuk angka nominal. Begitu juga jenis pengukuran yang digunakan, yaitu skala Gutmann. Dengan skala Gutmann, variabel yang diukur dijabarkan menjadi indikator variabel. Kemudian indikator tersebut dijadikan titik tolak untuk menyusun instrumen yang berupa pernyataan, nilai *ratio* dikonversi dari hasil prediksi berupa data nominal, nilai satu untuk hasil yang benar pada prediksinya ketika hasil prediksi sesuai dengan suara *input*, dan nilai nol pada prediksi yang salah ketika hasil prediksi tidak sama dengan suara *input*.

Beberapa variabel yang menjadi instrumen pada penelitian yaitu

- 3.4.1 variabel suara *input* yang sekaligus menjadi indikator bagi variabel *input* itu sendiri,

- 3.4.2 variabel fitur menjadi indikator dirinya sendiri yang bisa diukur dengan besaran nilai ekstraksi,
- 3.4.3 variabel jenis *learning* menjadi indikator bagi variabel jenis *learning*,
- 3.4.4 variabel kondisi mempunyai indikator jarak sumber suara dan dan tingkat kebisingan,
- 3.4.5 variabel *output* yang menjadi variabel dependen mempunyai indikator suara *input*, fitur, kondisi pengujian, dan jenis *learning* agar bisa terwakili.

Mengenai teknik pengolahan data, peneliti menggunakan statistik deskriptif. Peneliti menggambarkan data yang telah dikumpulkan tanpa membuat kesimpulan yang bersifat umum/generalisasi. Peneliti memakai statistik deskriptif karena memang hanya mendeskripsikan data dari sampel dengan membuat table, grafik, dan membandingkan hasil dari beberapa data sampel, dan tidak membuat generalisasi, hanya membuat penyajian data dari sampel yang telah diolah.

Peneliti mengumpulkan data-data hasil ukur mengenai tiap variabel, membuat statistik termasuk berbagai macam penyajian data dan memaparkan hasil data sampel suara yang diujikan seperti tabel/diagram *output*, analisa data, dan sebagainya. Peneliti hanya membuat penyajian data dari sampel yang diambil, bukan dari segala sumber yang ada di lingkungan.

3.5 Perancangan

Rancangan pengujian melibatkan beberapa hal, diantaranya penggunaan *database* sebagai data pembanding, dan alur bagaimana pengujian dilakukan/berjalan.

3.5.1 Database Suara

Database suara yang dilibatkan adalah empat jenis suara yang bersifat *indoor*/umum dijumpai di lingkungan rumah bagian dalam. Jenis suara diperlakukan sebagai kelas dalam *database*. Dalam *database indoor* terdiri dari 4 kelas sebagai berikut.

- Suara Bayi menangis/*baby_crying* sebanyak 100 *record*.
- Suara bel pintu/*door_bell* sebanyak 100 *record*.
- Suara ketukan pintu/*knocking_door* sebanyak 100 *record*.

- Suara air mengalir/*water_runs* sebanyak 100 *record*.

Tiap *record* adalah file berbentuk *wav*, berukuran panjang 5 detik. Pada tiap fitur mempunyai *database*-nya sendiri. Hal ini dikarenakan tiap fitur mengekstrak nilai-nilai yang berbeda satu sama lain.

3.5.2 System Flow Chart Pengujian



Gambar 3.1 System Flow Chart Pengujian.

Sebelumnya, *database* sudah di-load dalam *workspace* MATLAB. Pada Gambar 3.1, adalah diagram *sytem flow chart* bagaimana proses pengujian dilakukan. Tahap pertama, dilakukan perekaman suara selama lima detik dengan kondisi yang telah disiapkan. Suara ini yang dicari tahu prediksinya. Tahap kedua suara yang direkam diekstrak dengan fitur masing-masing. Tahap ketiga, hasil ekstraksi dibandingkan dengan *database* dengan fitur ekstraksi

yang sama. Cara pembandingannya dengan menggunakan jenis *learning*. Tahap keempat adalah menampilkan hasil prediksi suara yang sebelumnya direkam.

Penjelasan sebelumnya berlaku untuk satu kali prediksi. Dalam pengujian, dilakukan sepuluh kali pengujian pada tiap suara, pada tiap kondisi, pada tiap fitur, dan pada tiap jenis *learning*. Pada akhirnya dilakukan empat ribu kali pengujian pada penelitian ini, untuk mengetahui kriteria optimalisasi *supervised learning*. Tujuan pengujian pada penelitian ini adalah mencari tahu kriteria mana yang menghasilkan jumlah keakuratan/akurasi yang paling tinggi pada proses prediksi.

Akurasi dianggap bisa mewakili performa dalam aplikasi AudiTion. Jika hasil pengujian masih tidak bisa menghasilkan performa/akurasi yang tinggi sebesar 90%, akan dilakukan beberapa pengujian lanjutan. Pengujian lanjutan dilakukan pada kriteria fitur, jenis *learning*, dan kondisi dengan persentase tertinggi pada pengujian. Bentuk pengujian lanjutan berupa perubahan *database*. Pertama, *database* dikurangi sebanyak setengah atau lima puluh persen pada tiap suara. Kedua, pengurangan jenis suara dari empat menjadi tiga jenis suara saja. Ketiga, lama perekaman suara dikurangi dari lima detik menjadi satu detik saja. Terakhir, hasil pengujian lanjutan akan dibandingkan dengan menggunakan jenis suara lain, yaitu jenis suara *outdoor*.

Pengujian dilakukan pada 5 jenis fitur, yaitu

- MFCC,
- *centroid*,
- RMS,
- *low energy*, dan
- kombinasi gabungan *centroid*, RMS, & *low energy*.

Kedua, pengujian dilakukan pada jenis *learning*

- *discriminant analysis*,
- *AdaBoostM2*,

- KNN,
- *naive bayes*, dan
- *decision tree*.

Ketiga, pengujian juga dilakukan pada empat kondisi pengujian, yaitu

- kondisi sumber suara dekat dengan jarak interval 10-20cm & situasi hening dimana *noise*-nya sangat rendah/nyaris tidak ada serta tanpa adanya suara latar belakang.
- kondisi sumber suara dekat & situasi berisik/bising dimana ditambahkan *noise* saat perekaman berupa suara acara televisi, kipas angin, dll.,
- kondisi sumber suara jauh dengan jarak interval 1 meter & situasi hening, dan
- kondisi sumber suara jauh & situasi bising.