



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:

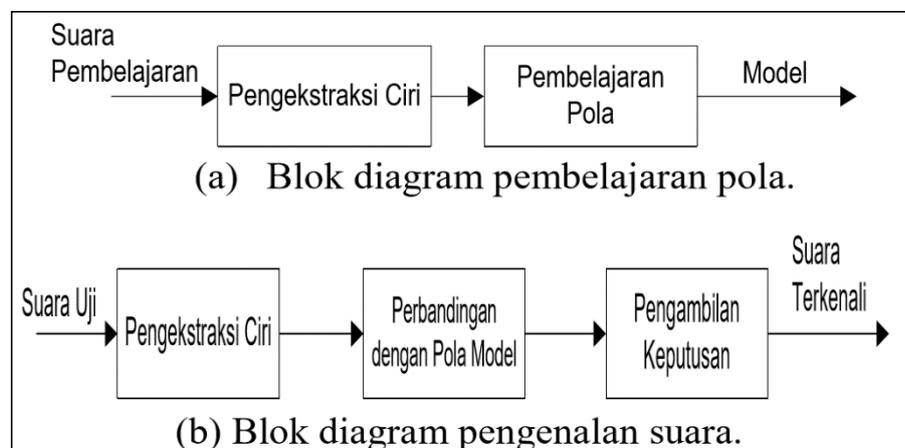
This license lets you remix, tweak, and build upon work non-commercially, as long as you credit the origin creator and license it on your new creations under the identical terms.

BAB II

LANDASAN TEORI

2.1 Pengenalan Suara

Pengenalan pola merupakan salah satu pendekatan pada pengenalan suara yang terdiri dari 2 langkah yaitu pembelajaran pola suara dan pengenalan suara melalui perbandingan pola. Tahap perbandingan pola adalah tahap saat suara yang dikenali dibandingkan polanya dengan setiap kemungkinan pola yang telah dipelajari dalam fase pembelajaran (Irfandy, 2010). Blok diagram pembelajaran pola dan pengenalan suara ditunjukkan pada Gambar 2.1.



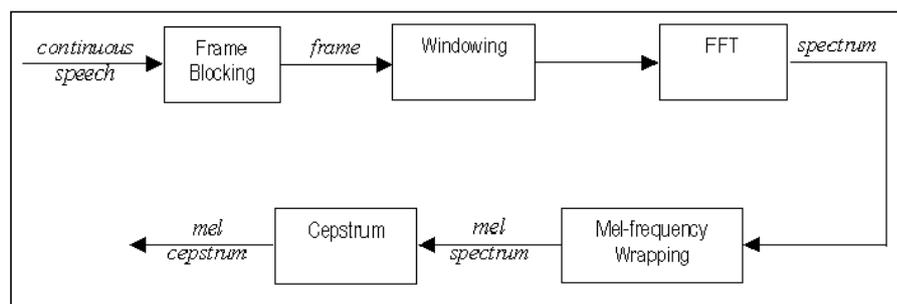
Gambar 2.1 Blok diagram pembelajaran pola dan pengenalan suara (Irfandy, 2010)

Pengenalan suara dibagi menjadi empat tahap, yaitu ekstraksi ciri yang merupakan proses untuk mendapatkan sederetan besaran pada bagian sinyal yang dimasukan untuk menetapkan pola pembelajaran. Kemudian pembelajaran pola yang digunakan untuk membuat pola representatif dari ciri-ciri suara yang kemudian menjadi pola referensi untuk pembanding dengan suara yang dimasukan. Tahap perbandingan dengan pola model dimana pola suara yang diberikan secara *real time* dibandingkan dengan pola referensi yang dihitung setiap kesamaan besar

dari pola suara yang ada dan tahap selanjutnya adalah tahap pengambilan keputusan dimana pada tahap ini merupakan tahap menentukan pola suara mana yang paling cocok untuk pola uji berdasarkan klasifikasi pola (Setiawan, dkk., 2011)

2.2 Mel-Frequency Cepstrum Coefficient (MFCC)

Mel-Frequency Cepstrum Coefficient (MFCC) bekerja berdasarkan prinsip dari sistem pendengaran manusia dalam skala mel, sehingga karakteristik parameter sinyal input yang dipilih sesuai dengan parameter yang diinginkan. Metode ini dapat mampu menangkap karakteristik suara dan dapat menghasilkan data seminimal mungkin tanpa menghilangkan informasi yang penting serta mereplikasikan organ pendengaran manusia dalam melakukan persepsi terhadap sinyal suara (Amalia, 2017). Tahapan untuk mendapatkan suatu *mel cepstrum* ditunjukkan pada Gambar 2.2.



Gambar 2.2 Diagram blok pemroses MFCC (Prasetyawan, 2016)

2.2.1 *Frame Blocking*

Pada proses ini sinyal suara disegmentasi menjadi beberapa *frame* yang saling tumpang tindih (*overlap*). Proses ini terus berlanjut sampai seluruh sinyal sudah masuk ke dalam satu atau lebih *frame*. Segmentasi *frame* yang *overlap* dilakukan agar tidak ada sinyal yang hilang. Proses ini menghasilkan matriks dengan N baris dan beberapa kolom sinyal $X[N]$ (Setiawan, dkk., 2011).

2.2.2 Windowing

Windowing setiap *frame* dilakukan untuk meminimalisir diskontinuitas sinyal pada permulaan dan akhir setiap *frame* (Setiawan, dkk., 2011). Bila *window* didefinisikan sebagai $w(n)$, $0 \leq n \leq N-1$, dimana N adalah jumlah sampel dalam tiap *frame*.

2.2.3 Fast Fourier Transform (FFT)

Fast Fourier Transform (FFT) adalah algoritma cepat dari *Discrete Fourier Transform* (DFT) untuk mengkonversi *frame* yang berisi N sampel dari domain waktu ke dalam domain frekuensi. Hasil dari proses ini disebut dengan *spectrum* (W.S & Salleh, 2014).

2.2.4 Mel-Frequency Wrapping

Persepsi manusia tentang frekuensi suara untuk sinyal ucapan tidak mengikuti skala linear. Untuk setiap nada dengan frekuensi sesungguhnya f , diukur dalam Hz, sebuah pola diukur dalam sebuah skala yang disebut 'mel'. Skala *mel-frequency* merupakan skala frekuensi rendah yang bersifat linear di bawah 1000 Hz dan skala frekuensi tinggi yang bersifat logaritmik di atas 1000 Hz (W.S & Salleh, 2014).

2.2.5 Cepstrum

Cepstrum biasa digunakan untuk mendapatkan informasi dari satu sinyal suara yang diucapkan oleh manusia. *Spectrum log mel* dikonversi menjadi *cepstrum* menggunakan *Discrete Cosine Transform* (DCT). Hasil dari proses ini di namakan *Mel-Frequency Cepstrum Coefficients* (MFCC) (Setiawan, dkk., 2011).

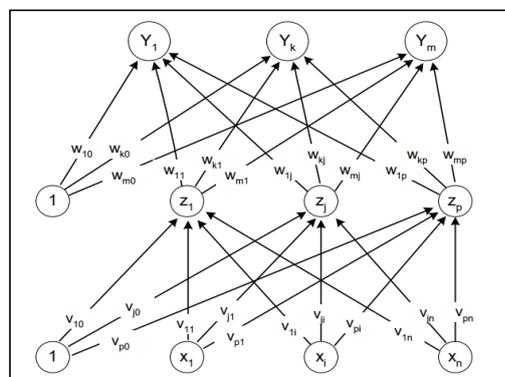
2.3 Jaringan Syaraf Tiruan

Jaringan syaraf tiruan didefinisikan sebagai sistem pemrosesan informasi yang menyerupai jaringan syaraf manusia (Amalia, 2017). Jaringan syaraf manusia didasari dari pemahaman manusia tentang generalisasi model matematis dengan pemrosesan informasi yang terjadi pada *neuron* yang dihubungkan satu sama lain. sinyal mengalir dari satu *neuron* ke *neuron* lain melalui sambungan ini dan setiap sambungan memiliki bobot yang digunakan untuk menggandakan sinyal yang dikirim, setiap neuron melakukan aktifasi terhadap sinyal yang masuk sebagai sinyal keluarannya (Amalia, 2017).

2.4 Backpropagation

2.4.1 Standar Backpropagation

Jaringan syaraf tiruan dengan menggunakan lapis tunggal memiliki keterbatasan dalam pengenalan pola, dengan menambahkan beberapa lapisan tersembunyi maka kelemahan ini dapat ditanggulangi. Pelatihan dengan menggunakan beberapa lapisan memakan waktu yang lama (Siang, 2009).



Gambar 2.3 Arsitektur Backpropagation (Siang, 2009)

Gambar 6.4 merupakan arsitektur *backpropagation* dengan n buah masukan (x_1, \dots, x_n) dengan penambahan sebuah bias, sebuah layar tersembunyi yang terdiri

dari p unit (z_1, \dots, z_p) dengan penambahan sebuah bias, serta m buah unit keluaran (Y_1, \dots, Y_m) , Dimana v_{ji} merupakan bobot garis dari unit masukan x_i ke unit layar tersembunyi z_j . v_{j0} menyatakan bobot garis yang menghubungkan bias di unit masukan ke unit layar tersembunyi z_j . w_{kj} merupakan bobot dari unit layar tersembunyi z_j ke unit keluaran y_k . w_{k0} menyatakan bobot dari bias di layar tersembunyi ke unit keluaran z_k (Siang, 2009).

2.4.2 Pelatihan Standar Backpropagation

Pelatihan dilakukan melalui tiga fase, fase yang pertama adalah fase maju dimana pola masukan dihitung maju mulai dari *input* sampai *output* menggunakan fungsi aktivasi yang ditentukan. Rumus yang digunakan untuk mencari nilai masukan pada fase maju adalah sebagai berikut.

$$z_{net_j} = v_j + \sum_{i=1}^n x_i v_{ji} \quad \dots(2.1)$$

Fase kedua adalah fase mundur dimana pada fase ini selisih antara *output* jaringan dengan target yang diinginkan merupakan kesalahan yang terjadi. Rumus yang digunakan untuk menghitung faktor kesalahan yang terjadi adalah sebagai berikut.

$$\delta_k = (t_k - y_k) f'(y_{net_k}) = (t_k - y_k) y_k (1 - y_k) \quad \dots(2.2)$$

Fase ketiga adalah modifikasi bobot untuk menurunkan kesalahan yang terjadi (Siang, 2009). Pada fase modifikasi nilai bobot yang dibawa dimodifikasi agar mengurangi terjadinya kesalahan dengan menggunakan rumus sebagai berikut.

$$w_{kj}(\text{baru}) = w_{kj}(\text{lama}) + \Delta w_{kj} \quad \dots(2.3)$$

Semua fase dilakukan secara berulang secara terus menerus sampai mencapai kondisi yang sudah ditentukan sebelumnya terpenuhi yaitu melebihi jumlah maksimum iterasi ataupun melebihi batas toleransi yang ditentukan (Putra, dkk.,

2007). Kemudian akan dilakukan penghitungan persentase tingkat akurasi yang dihitung dengan persamaan sebagai berikut.

$$\text{Persentase} = \frac{\text{jumlah data yang dikenali}}{\text{jumlah data uji coba}} \times 100\% \quad \dots(2.4)$$

2.4.3 Fungsi Aktivasi

Dalam backpropagation fungsi aktivasi yang dipakai harus memenuhi beberapa syarat yaitu kontinu, terdiferensial dengan mudah dan merupakan fungsi yang tidak turun. Salah satu fungsi yang memenuhi ketiga syarat tersebut adalah fungsi sigmoid biner yang memiliki *range* (0,1). Rumus yang digunakan untuk fungsi sigmoid biner adalah sebagai berikut (Siang, 2009).

$$f(x) = \frac{1}{1+e^{-x}} \quad \dots(2.4)$$

Dengan turunan :

$$f'(x) = f(x)(1 - f(x)) \quad \dots(2.5)$$

Fungsi lain yang sering dipakai adalah fungsi sigmoid bipolar yang bentuk fungsinya mirip dengan fungsi sigmoid biner, tetapi memiliki *range* (-1,1). Rumus yang digunakan untuk fungsi sigmoid bipolar adalah sebagai berikut.

$$f(x) = \frac{1}{1+e^{-x}} - 1 \quad \dots(2.6)$$

Dengan turunan :

$$f'(x) = \frac{(1+f(x))(1-f(x))}{2} \quad \dots(2.7)$$

Fungsi sigmoid memiliki nilai maksimum =1. Maka untuk pola yang targetnya > 1, pola masukan dan keluaran harus terlebih dahulu ditransformasi sehingga semua polanya memiliki rentang yang sama seperti fungsi sigmoid yang dipakai. Alternatif lain adalah menggunakan fungsi aktivasi sigmoid hanya pada

layar yang bukan layar keluaran. Pada layar keluaran, fungsi aktivasi yang dipakai adalah fungsi identitas yakni $f x = x$ (Siang, 2009).

2.5 Arduino Uno

Arduino Uno merupakan *platform* pembuatan *prototype* elektronik yang bersifat *open-source hardware* yang berdasarkan pada perangkat keras dan perangkat lunak yang fleksibel dan mudah digunakan (Tambak & Bahriun, 2015). Arduino Uno memiliki 14 *digital input/output pin*, 6 *input analog* (6 pin untuk *output PWM*), 16 MHz *osilator* kristal, koneksi USB, jack listrik tombol reset. Pin-pin yang ada pada Arduino Uno diperlukan untuk mendukung mikrokontroler, hanya terhubung ke komputer dengan menggunakan kabel USB atau sumber tegangan dari adaptor AC-DC atau baterai untuk menggunakan Arduino Uno (Saputri, dkk., 2014).