

UTS PENGENALAN POLA



DISUSUN OLEH:

**Frelly Lumba – F55118282
Sean Demas – F55118286
Cindy Noviandini – F55118255**

KELAS C

**PROGRAM STUDI TEKNIK INFORMATIKA
JURUSAN TEKNOLOGI INFORMASI
FAKULTAS TEKNIK
UNIVERSITAS TADULAKO
2022**

1. Data klasifikasi penyakit tuberklosis(TB)

Tuberkulosis (TB) sebuah penyakit yang dipicu oleh kuman yang memiliki nama *Mycrobacterium tuberculosis*, penyakit TB termasuk kedalam penyakit infeksi menular. Kuman *Mycrobacterium* memiliki beberapa jenis spesies diantaranya adalah *M.tuberculosis*, *M.bovis*, *M.Leprae*, *M.africanum*, dll (Kementrian Kesehatan Republik Indonesia, 2018). Gejala yang paling utama pada pengidap TB adalah batuk selama 2 minggu ataupun lebih,

Gejala batuk ini biasanya juga diikuti dengan gejala lainnya seperti batuk berdarah dan berdahak, mengalami sesak nafas, badan akan menjadi lebih mudah lelah dan lemas, tiap malam hari badan akan mudah berkeringat, serta penderita akan mengalami penurunan nafsu makan. Penyakit TB akan menyerang paru-paru dan jika tidak mendapat pengobatan akan menyebar ke organ tubuh lainnya. Menurut World Health Organization (WHO) (2018), kasus TB di Indonesia mengakibatkan 301 orang meninggal akibat penyakit ini.

2. Metode yg di gunakan Metode Extreme Learning Machine (ELM)

ELM dipopulerkan Guang-Bin Huang, metode ini memiliki arsitektur berupa Single Layer Feedforward Networks (SLFNs) yakni satu input layer, satu hidden layer, dan satu output layer. ELM mempunyai kecepatan yang lebih baik dalam learning dibanding algoritme lainnya dan menghasilkan generalisasi yang baik sehingga menghasilkan error cenderung kecil (Huang, Zhu, & Siew, 2004). Penetapan bobot maupun bias pada metode ELM dikerjakan secara acak dalam waktu bersamaan. Hal ini akan mempercepat proses pembelajaran dan akan membuat hasil prediksi.

2.1. Normalisasi Data

Metode yang dipakai yakni Min-Max Normalization.

$$d^- = \frac{d-min}{max-min}$$

Keterangan :

d^- = Hasil normalisasi.

d = Nilai rata-rata

min = min. data

max = max.data

2.2. Training Data

- a. Inialisasi bobot dengan rentang $[-0,5,0,5]$ serta bias dengan rentang $[0,1]$ secara acak.
- b. Perhitungan output hidden layer menggunakan Persamaan berikut;
$$H_{init} = X \cdot WT + \text{ones}(N_{train}, 1) \cdot b$$

Keterangan:

H_{init} Hasil perkalian dari matriks X dan WT

X = Matriks data.

Wt = Matriks transposisi bobot.

b = Matriks bias.

- c. Perhitungan H menggunakan sigmoid biner dengan Persamaan berikut:

$$H = 1 + \frac{1}{1 + e^{-(H_{init})}}$$

Keterangan:

H = Matriks H.

e = Bilangan euler pembulatan bernilai 2,718.

H_{init} = Hasil perkalian dari matriks X dan Wt .

- d. Perhitungan Moore-Penrose pseudo inverse dengan Persamaan berikut:

$$H^{+} = (H(x) T \cdot H(x))^{-1} \cdot H T$$

Keterangan:

H^{+} = Hasil matriks.

$(H(x) T \cdot H(x))^{-1}$ = Matriks inverse perkalian H dan H Transpose.

- e. Perhitungan output weight dengan Persamaan berikut:

$$\beta = H^{+} \cdot T$$

Keterangan:

β = Hasil output weight.

H^{+} = Matriks Moore-Penrose pseudo inverse

- f. Perhitungan output layer menggunakan Persamaan berikut:

$$Y = H(x) \cdot \beta$$

Keterangan:

Y = Hasil matriks.

H = Output hidden layer.

β = Output weight.

2.3. Testing Data

1. Bobot dan bias secara random yang diperoleh dari training.
2. Perhitungan matriks Hinit menggunakan Persamaan berikut

$$H_{init} = X \cdot Wt + \text{ones}(N_{test}, 1) \cdot b$$

Keterangan:

H_{init} = Matriks hasil.

X = Matriks data.

Wt = Matriks transposisi bobot.

b = Matriks bias.

3. Perhitungan matriks H menggunakan sigmoid biner dengan Persamaan berikut:

$$H = 1 + \frac{1}{1 + e^{-(h_{init})}}$$

Keterangan:

H = Matriks H.

e = Bilangan euler pembulatan bernilai 2,718.

H_{init} = Hasil perkalian dari matriks X dan Wt .

4. Perhitungan matriks output layer dengan Persamaan 12, nilai output weight dari proses training.

$$Y = H(x) \cdot \beta$$

Keterangan:

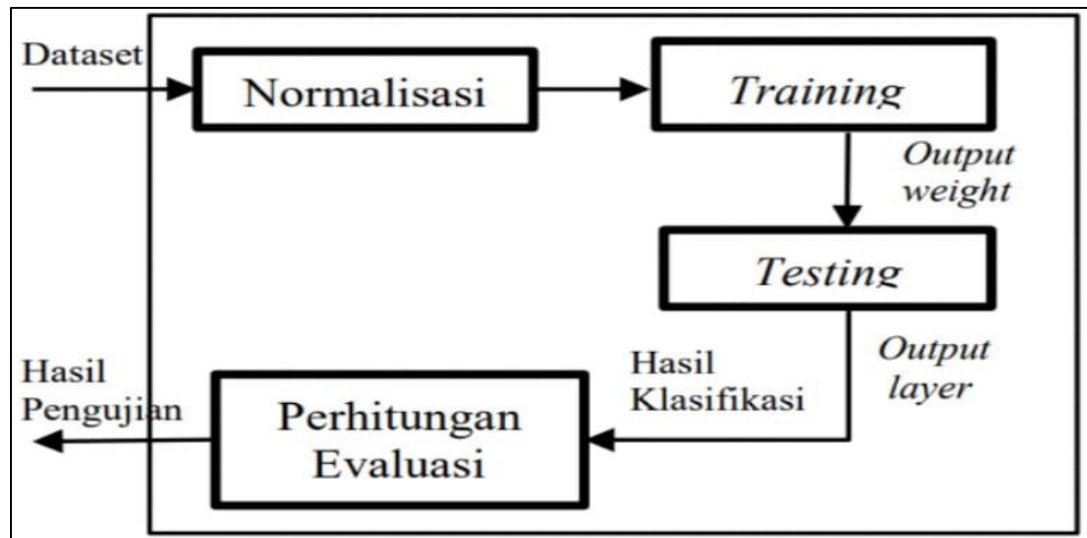
Y = Hasil output layer.

H = Matriks H

β = Matriks output weight.

3. METODOLOGI PENELITIAN

Dalam penelitian ini memakai metode ELM untuk klasifikasi penyakit TB. Klasifikasi TB terdapat 2 kelas yakni kelas 1 adalah kelas positif penyakit TB dan kelas 2 adalah negatif penyakit TB. Metode ELM terdapat tahapan-tahapan yang harus diselesaikan secara urut. Dataset akan dinormalisasi, pada tahap training akan menghasilkan nilai output weight nantinya nilai tersebut dipakai dalam proses testing untuk menghasilkan nilai output layer. Selanjutnya setelah mengetahui hasil klasifikasi dilakukan proses perhitungan pengujian menggunakan confusion matrix dengan parameter sensitivity, specificity, dan accuracy untuk menghitung seberapa baik kerja sistem dalam melakukan proses klasifikasi. Gambar 2 diagram klasifikasi menggunakan metode ELM.



Teknik pengumpulan data memakai wawancara dan melakukan observasi terhadap kasus TB di Puskesmas Dinoyo Kota Malang. Data yang dipakai yakni data primer yang langsung diolah dan diperoleh melalui Puskesmas Dinoyo Kota Malang. Data tersebut adalah data penyakit TB tahun 2018-2019 Puskesmas Dinoyo sebanyak 100 data dengan 24 fitur dan terdapat 2 kelas yakni kelas 1 merupakan pasien positif penyakit TB dan kelas 2 merupakan pasien negatif terhadap penyakit TB. Data TB diolah untuk proses klasifikasi dan dilakukan proses pembagian data yang terbagi menjadi data training dan data testing.

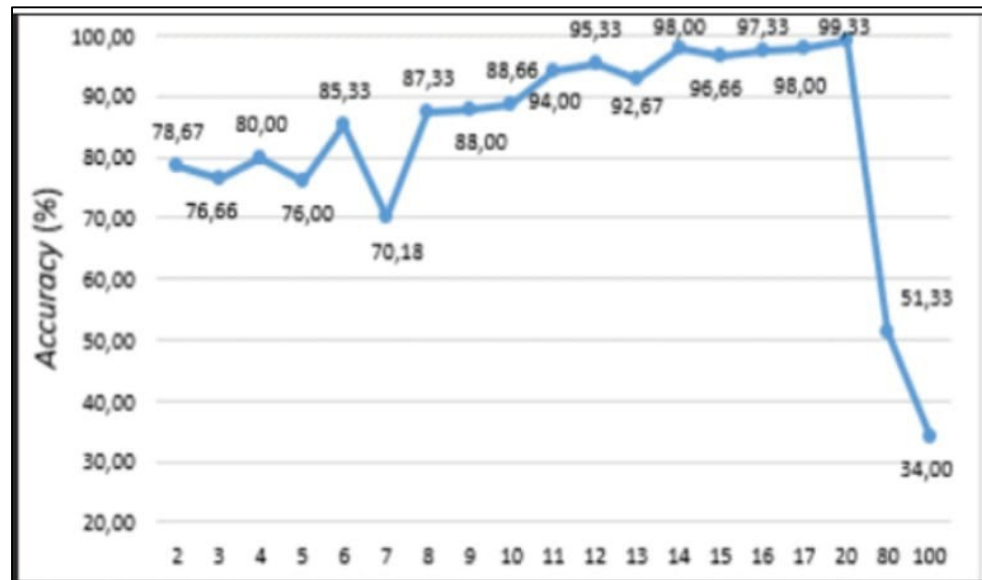
4. PENGUJIAN DAN ANALISIS

Penelitian klasifikasi penyakit Tuberkulosis (TB) menggunakan metode Extreme Learning Machine (ELM) menggunakan beberapa pengujian untuk mendapatkan parameter optimal. Dengan parameter optimal maka penelitian akan memperoleh hasil yang bagus pada sistem. Penelitian ini menggunakan 3 macam pengujian. Pengujian tersebut adalah pengujian jumlah hidden neuron, pemakaian fungsi aktivasi, serta pengaruh jumlah data training dengan data testing. Data yang digunakan mempunyai 2 kelas yakni kelas 1 dan kelas 2. Pada kelas 1 merupakan pasien positif penyakit TB dan pada kelas 2 merupakan pasien negatif terhadap penyakit TB. Tiap pengujian terdapat proses perhitungan sensitivity, specificity, dan accuracy.

4.1. Pengujian Pengaruh Jumlah Hidden Neuron

Tujuan pengujian ini untuk melihat pengaruh dari banyak jumlah hidden neuron pada hasil klasifikasi dan tingkat akurasi. Persentase data training dan data testing yang digunakan 70%:30% yakni 70 data training dan 30 data testing, dengan aktivasi sigmoid biner. Pengujian hidden neuron dilakukan pada jumlah 2-20, 80, dan 100. Setiap hidden neuron dilakukan 5 kali pengujian. Hasil pengujian berbeda-beda dikarenakan pada tiap pengujian data yang diolah akan dilakukan

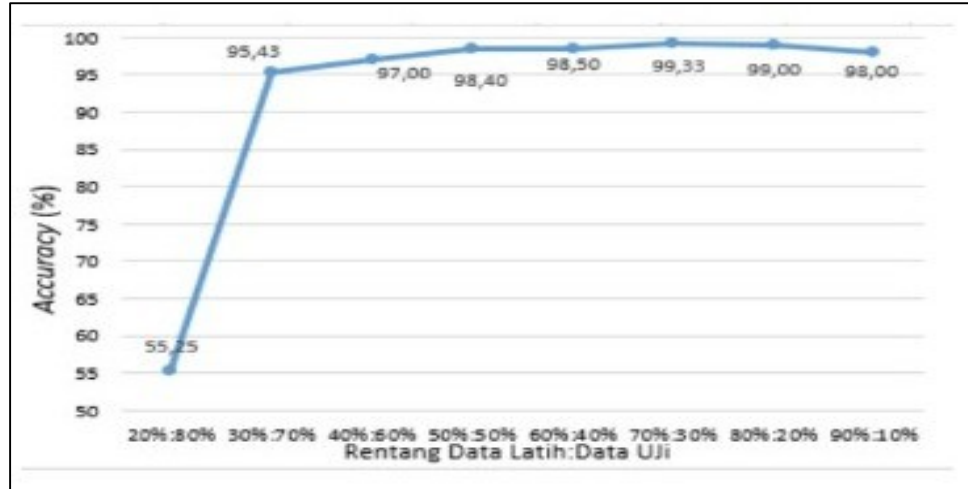
random, nilai bobot dan bias juga diinisialisasi dengan random. seperti gambar dibawah ini



Nilai dari rata-rata pengujian jumlah hidden neuron yang paling baik pada Gambar 3 adalah pada jumlah hidden neuron sebanyak 20 dengan hasil rata-rata sensitivity 1,00, specificity 0,98, dan accuracy 99,33%. Pada hidden neuron 2 sampai 20 menghasilkan nilai sensitivity, specificity, dan accuracy semakin meningkat. Peningkatan terjadi signifikan hal ini menunjukkan bahwa kelas 1 maupun kelas 2 pada data TB terklasifikasi dengan baik dan benar. Semakin banyak jumlah hidden neuron juga mempengaruhi besar matriks terhadap ruang fitur dalam melakukan proses learning. Akan tetapi pada jumlah hidden neuron 80 dan 100 nilai sensitivity, specificity, dan accuracy mengalami penurunan drastis. Hal ini disebabkan dengan menggunakan hidden neuron terlalu banyak akan membuat proses kurang optimal. Jumlah hidden neuron yang baik dan menghasilkan nilai akurasi tinggi terdapat pada rentang 20-60 (Xiao, Li, & Mao, 2017).

4.2. Training dan Data Testing

Pengujian ini menggunakan metode holdout dengan melakukan pembagian data training dan data testing menggunakan 20%:80% hingga 90%:10%. Digunakan metode holdout dikarenakan jumlah data yang kurang memadai yang hanya berjumlah 100 dan holdout lebih mudah dalam penerapan. Pengujian pengaruh data training dengan data testing bertujuan untuk mengetahui bagaimana pengaruh rentang data training dan data testing terhadap hasil nilai akurasi. Pengujian ini menggunakan jumlah hidden neuron sebanyak 20 dan fungsi aktivasi sigmoid biner, tiap rentang dilakukan 5 kali pengujian. Gambar 5 merupakan grafik hasil pengujian pengaruh jumlah Data training dan Data testing



Nilai dari rata-rata pengujian pengaruh jumlah data training dan data testing yang terbaik dari Gambar 5 adalah pada persentase 70% data training dan 30% data testing dengan menghasilkan nilai sensitivity 0,99, specificity 1,00, dan accuracy 99,33. Pada persentase 20%:80% sampai 60%:40% menghasilkan nilai sensitivity, specificity, dan accuracy semakin meningkat. Pada persentase 70%:30% sampai 90%:10% mengalami hanya sedikit penurunan pada sensitivity, specificity, dan accuracy. Hal ini menunjukkan bahwa data yang digunakan dapat mengikuti pola data dan menghasilkan akurasi yang tinggi.