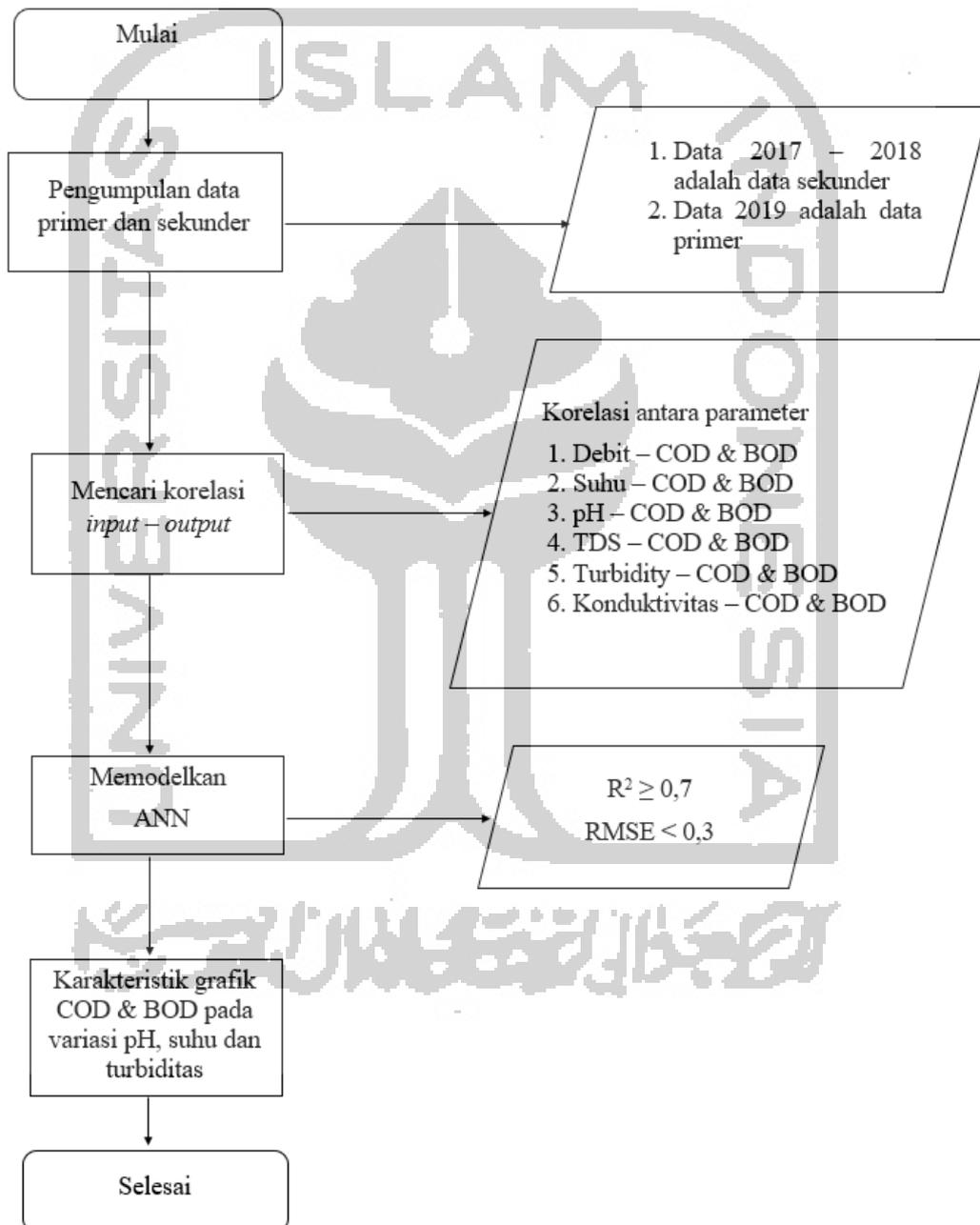


BAB III METODE PENELITIAN

3.1 Tahapan Penelitian

Tahapan penelitian yang dilaksanakan tergambar pada Gambar 3.1 berikut

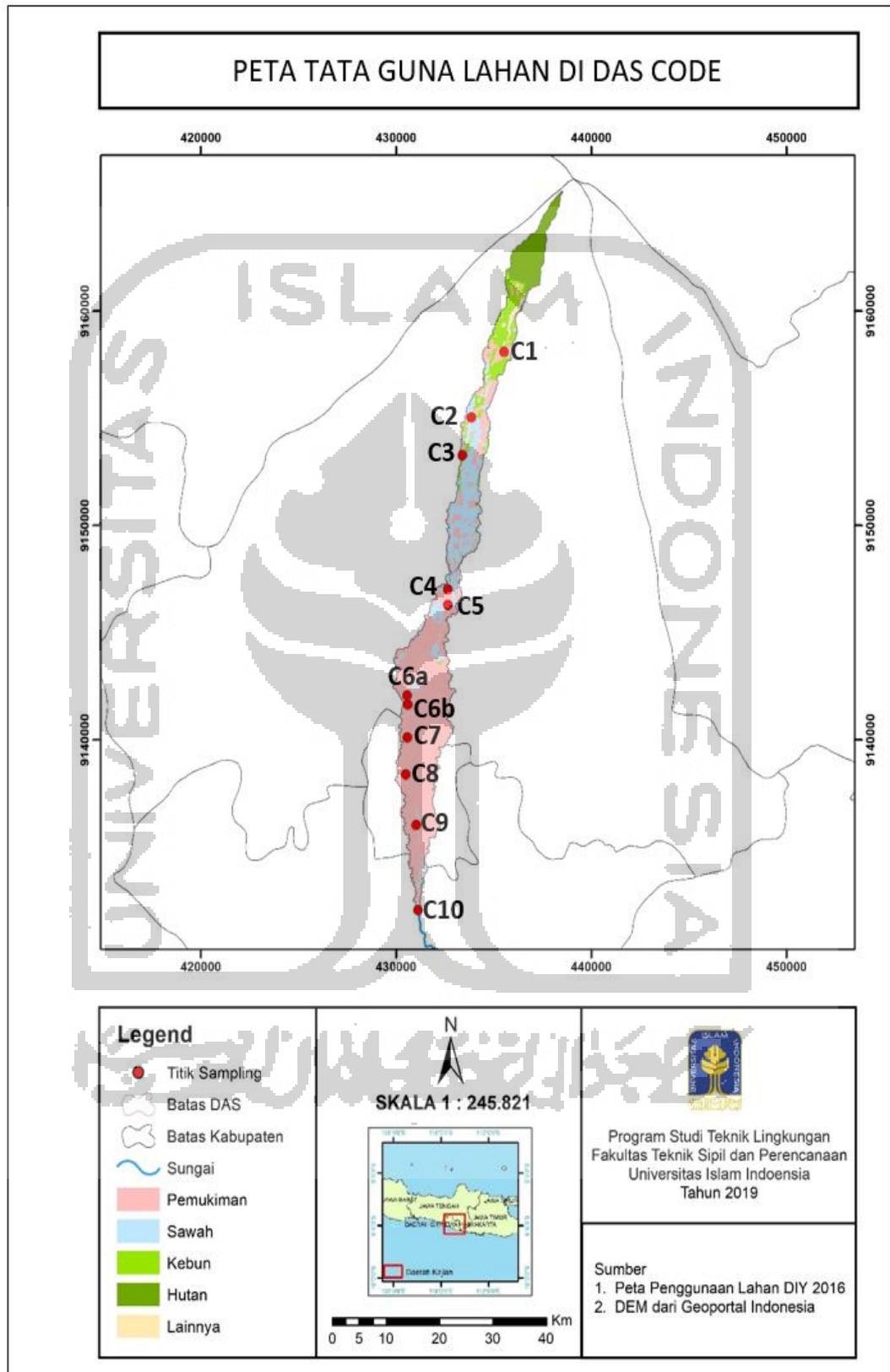


Gambar 3.1 Tahapan Penelitian

3.2 Waktu dan Lokasi Penelitian

Penelitian ini dilaksanakan mulai April 2019 hingga Juli 2019. Pelaksanaan penelitian terbagi menjadi dua proses yaitu kegiatan pengumpulan data dan membangun model ANN. Pengumpulan data dimulai dengan sampling dan pengujian laboratorium. Sampling bulan pertama hingga bulan keempat dilaksanakan pada tanggal 25 April 2019, 16 Mei 2019, 27 Juni 2019 dan 25 Juli 2019, sedangkan pengujian laboratorium dilakukan pada hari berikutnya setelah pengambilan sampel.

Lokasi pengambilan sampel dari hulu ke hilir berada pada daerah yang memiliki peruntukan lahan dan kondisi yang berbeda-beda seperti yang tertera dalam Tabel 3.1. Berdasarkan peta tata guna lahan DIY 2016 daerah hulu dijadikan sebagai kawasan hutan dan perkebunan, sedangkan semakin ke hilir tata guna lahan berubah fungsi menjadi kawasan permukiman seperti yang ditunjukkan pada Gambar 3.2. Berdasarkan SNI 6989.57:2008 lokasi pemantauan kualitas air dilakukan pada sumber air alamiah, sumber air tercemar, dan sumber air yang dimanfaatkan. Meninjau SNI 03-7016-2004, penentuan lokasi pengambilan sampel di daerah hulu sebagai *base line station* dan daerah hilir sebagai *impact station*. Merujuk pada SNI 03-7016-2004, lokasi pengambilan sampel secara keseluruhan dekat dengan jembatan. Lokasi jembatan memberikan keuntungan bagi peneliti yaitu mempermudah proses pengambilan sampel dan identifikasi pengambilan contoh lebih pasti.



Gambar 3.2 Peta Lokasi Pengambilan Sampel

(Raziq, 2018)

Tabel 3.1 Lokasi Pengambilan Sampel

No	Kode Sampel	Kondisi	Lokasi
1	C1	Permukiman, hutan	-7.614893 , 110.415769
2	C2	Permukiman, sawah irigasi, perkebunan	-7.6435716 , 110.3982948
3	C3	Permukiman, sawah irigasi, perkebunan	-7.659587 , 110.3940582
4	C4	Permukiman, sawah irigasi, perkebunan, (industri)	-7.7192692 , 110.3862587
5	C5	Permukiman, sawah irigasi	-7.7225154 , 110.3863138
6	C6A	Permukiman	-7.7623688 , 110.3684764
7	C6B	Permukiman	-7.7623688 , 110.3684764
8	C7	Permukiman	-7.7784614 , 110.3682342
9	C8	Permukiman	-7.7941157 , 110.3675385
10	C9	Permukiman, sawah irigasi	-7.8155034 , 110.3724641
11	C10	Permukiman, sawah irigasi	-7.8513282 , 110.3736023

(Raziq, 2018)

3.3 Pengumpulan Database

Database merupakan data yang dikumpulkan dari November 2017 hingga November 2018 dan April 2019 hingga Juli 2019, kecuali data pada bulan April 2018 dan Juni 2018 tidak digunakan. Kumpulan data tersebut mewakili kualitas Sungai Code pada tahun 2017, 2018 dan 2019. Data dari November 2017 hingga November 2018 merupakan data sekunder dan data dari April 2019 hingga Juli 2019 sebagai data primer. Proses pengumpulan data primer dengan melakukan sampling dan pengujian sampel.

Berdasarkan SNI 03-7016-2004 sampel diambil dengan metode *grab sample* dan *composite samples*. Metode *grab samples* diterapkan di lokasi pertama (site 1) dan lokasi lainnya menerapkan metode *composite samples*. Penerapan metode *composite samples* dilakukan dengan jumlah titik yang berbeda, terdapat lokasi yang terbagi menjadi tiga hingga lima titik pengambilan. Perbedaan jumlah titik pengambilan sampel disesuaikan dengan lebar sungai yang berbeda-beda. Jumlah sampel yang diambil pada tiap titiknya sebanyak 1 jerigen dengan ukuran 2,5 liter. Kemudian air yang telah diambil dicampur dan dihomogenkan.

Parameter-parameter yang diuji terbagi menjadi dua parameter yakni parameter fisika dan kimia.

3.3.1 Parameter Fisika

Parameter fisika adalah parameter yang dapat diamati akibat terjadinya perubahan fisik air. Parameter-parameter fisika yang seringkali digunakan dalam penentuan kualitas air adalah cahaya, suhu, kecerahan, kekeruhan, warna, konduktivitas, padatan total, padatan terlarut, padatan tersuspensi dan salinitas (Effendi, 2012).

Parameter-parameter fisik yang diuji dalam penelitian ini meliputi pH, suhu, dan turbiditas. Keseluruhan parameter tersebut diuji langsung di lokasi pengambilan sampel. Metode yang digunakan dalam pengujian parameter-parameter fisik tersebut tertera dalam Tabel 3.2

Tabel 3.2 Alat dan Metode untuk Parameter Fisika

No	Parameter (<i>input</i>)	Alat/Metode	Unit
1	pH	pH meter dan pH universal	-
2	Suhu	Probe TDS	°C
3	Turbiditas	Turbidimeter	NTU

3.3.2 Parameter Kimia

Parameter kimia menjadi salah satu parameter yang dijadikan sebagai acuan dalam menentukan kualitas air. Penelitian ini melakukan pengujian terhadap dua parameter COD dan BOD. Kedua parameter tersebut diuji sesuai dengan metode yang sudah ditetapkan seperti yang tertera pada Tabel 3.3

Tabel 3.3 Alat dan Metode untuk Parameter Kimia

No	Parameter (<i>output</i>)	Alat/Metode	Unit	SNI
1	COD	Uji Spektrofotometri	mg/l	SNI 6989.2:2009
2	BOD	Uji Titrasi	mg/l	SNI 06-6989.14-2004

3.3.3 Database

Penelitian ini mengembangkan dua model dengan nilai bobot yang berbeda. Pengembangan model yang pertama untuk memprediksi COD dan yang kedua untuk memprediksi BOD, sehingga jumlah data yang digunakan sebagai dasar membangun model berbeda dengan menggunakan 142 data untuk COD dan 113 untuk BOD. Jumlah data yang digunakan dalam membangun model terhitung banyak, maka untuk mengetahui ketimpangan data dibutuhkan penentuan standar deviasi dengan persamaan berikut,

$$s = \sqrt{\frac{1}{n-1} \sum_i^n (x_i - \bar{x})^2} \quad (1)$$

Keterangan,

n = Jumlah data xi = Nilai data x = Rata-rata sampel

3.4 Pencarian Korelasi Data Output dengan Input

Database yang dijadikan sebagai data pembelajaran ANN terdiri dari lima parameter yang terbagi menjadi data *input* dan *output* dan jumlah penggunaan data hingga 142 dan 113. Pemodelan ANN membutuhkan analisa korelasi antara parameter. Tujuan dari analisa korelasi tersebut untuk mengetahui apakah hubungan antara parameter tersebut positif atau negatif.

Penentuan korelasi dalam penelitian ini dengan menggunakan fungsi linier. Persamaan linier yang digunakan seperti rumus (2),

$$y = ax + b \quad (2)$$

Nilai R^2 dari proses pengkorelasian *database* menunjukkan nilai yang sangat kecil. Korelasi COD dengan pH, suhu dan turbiditas menunjukkan besar R^2 0,0217; 0,0046 dan 0,0016. Besar R^2 yang ditunjukkan pada korelasi BOD dengan pH, suhu dan turbiditas adalah 0,0079; 0,0003 dan 0,1045. Berdasarkan nilai-nilai tersebut tingkat kelinieran antar parameter sangat kecil, akan tetapi melalui garis linier yang terbentuk dapat disimpulkan korelasi yang terjadi bersifat positif atau negatif.

3.5 Pemodelan ANN

Artificial Neural Network (ANN) adalah model yang meniru cara kerja jaringan biologis. Metode yang akan digunakan dalam penelitian ini adalah metode *back propagation* dengan menggunakan *software excel* dikarenakan metode *propagation* memiliki kemampuan dalam menangani masalah pengenalan pola-pola yang kompleks. Sifat metode *backpropagation* yang *supervised* dan multilayer, maka dibutuhkan data training dalam prosesnya dan terdapat minimal *input layer*, *hidden layer* dan *output layer*. Perubahan nilai bobot dalam arah mundur (*backward*) algoritma *backpropagation* menggunakan *error output*. Penentuan nilai *error* harus didahului dengan mengerjakan perambatan maju (*forward propagation*), pada saat *forward propagation* neuron-neuron akan diaktifkan dengan menggunakan fungsi aktivasi (Yuliandar dkk., 2016).

Terdapat beberapa tahapan dalam penggunaan ANN dengan metode *propagation* yaitu,

1. Normalisasi data

Proses untuk melakukan skala data sehingga suatu data berada dalam suatu rentang nilai tertentu yang dilakukan untuk masing-masing data latih dan data uji yang akan digunakan (Sabati dkk., 2014). Menurut (Fernanda, 2012) setiap unit *input* (x_i , $i = 1, 2, \dots, n$) bertugas untuk menerima sinyal *input* x_i dan menyebarkan ke semua unit pada *hidden layer*. Persamaan matematis perhitungan normalisasi data yang digunakan adalah persamaan (3),

$$x' = \left(\frac{X_i - X_{\min}}{X_{\max} - X_{\min}} \right) \quad (3)$$

2. Perhitungan di dalam *hidden layer*

Menurut (Fernanda, 2012) setiap unit *hidden* (z_j , $j = 1, 2, \dots, p$) bertugas untuk menjumlahkan bobot. Persamaan matematis perhitungan dalam *hidden layer* yang digunakan adalah persamaan (4),

$$H = \sum X' w + BS w \quad (4)$$

3. Penerapan fungsi aktivasi data *input*

Penerapan fungsi aktivasi untuk menghitung sinyal *output* kemudian dikirimkan menuju *output layer* (Fernanda, 2012). Penerapan fungsi aktivasi (H^*) menggunakan fungsi sigmoid biner pada persamaan (5) dan (6),

$$H^* = f(y) = \frac{1}{1 + e^{-y}} \quad (5)$$

$$y = \sum_{i=0}^n X_i W_i \quad (6)$$

4. Penjumlahan bobot dengan sinyal *output*

Setiap unit *output* (Y_k , $k = 1, \dots, m$) bertugas untuk menjumlahkan bobot dari sinyal *output* (Fernanda, 2012). Persamaan matematis yang digunakan terdapat pada rumus (7),

$$Y = \sum H^* w + BS w \quad (7)$$

5. Penerapan fungsi aktivasi

Menerapkan fungsi aktivasi untuk menghitung sinyal *output* (Fernanda, 2012). Penetapan sinyal *output* menggunakan persamaan (8),

$$Y^* = f(y) = \frac{1}{1 + e^{-y}} \quad (8)$$

6. Penentuan *Root mean square error* (RMSE) dan koefisien determinasi (R^2)

Persamaan untuk menentukan nilai RMSE dan R^2 menurut (Sarkar, 2015) terdapat pada persamaan (9) dan (10),

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Y_i \text{ prediksi} - Y_i \text{ aktual})^2}{n}} \quad (9)$$

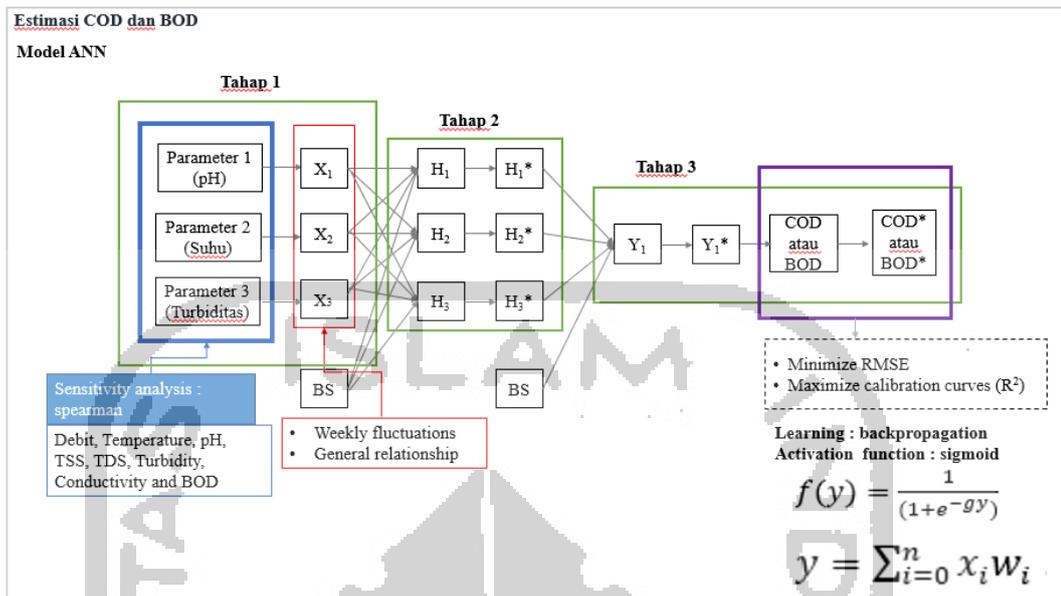
$$R^2 = 1 - \frac{\sum (Y_i \text{ prediksi} - Y_i \text{ aktual})^2}{\sum (Y_i \text{ prediksi} - Y_i \text{ rata-rata})^2} \quad (10)$$

Keterangan,

X	=	Unit <i>input</i>
X'	=	Normalisasi <i>input</i>
X*	=	Aktivasi <i>input</i>
w	=	Bobot
H	=	<i>Hidden layer</i>
H*	=	Hasil aktivasi H
Y'	=	Normalisasi <i>output</i>
Y	=	Unit <i>Output</i>
Y*	=	Hasil aktivasi Y
BS	=	Bias

RMSE = Root Mean Square Error

R^2 = Koefisien determinasi



Gambar 3.3 Mekanisme kerja ANN untuk memprediksi COD dan BOD

Berdasarkan Gambar 3.3 di atas rangkaian pemodelan ANN yang telah dihitung akan memiliki hasil akhir berupa nilai RMSE dan R^2 . Perhitungan dikatakan baik jika nilai RMSE mendekati 0 dan R^2 mendekati 1.

Rentang nilai RMSE yang digunakan adalah 1 artinya *error* tinggi, RMSE dengan *error* rendah terdeskripsikan nilai < 3 (Veerasamy dkk., 2011). Rentang nilai R^2 yang digunakan $< 0,3$ (sangat lemah); $0,3 < r < 0,5$ (rendah); $0,5 < r < 0,7$ (sedang) dan $r > 0,7$ (kuat) (Moore dkk., 2013).

Proses yang terjadi dalam pengembangan model ANN terbagi dua yaitu proses *training* dan proses *testing*. Proses *training* berfungsi untuk melatih jaringan saraf tiruan agar mampu melakukan identifikasi perubahan COD dan BOD (output) atas perubahan parameter input dengan terbentuknya arsitektur jaringan dan menetapkan nilai bobot untuk tahap pengenalan. Arsitektur jaringan yang telah diperoleh ditentukan oleh jumlah data training yang digunakan. Variabel training ditentukan oleh pengguna yang terdiri dari *hidden neuron*, *learning rate*, *epoch* dan *error goal*. Proses *testing* merupakan tahap pengujian model yang sudah dikembangkan dengan menggunakan arsitektur sebelumnya. Tingginya tingkat akurasi dipengaruhi oleh jumlah *hidden neuron*, 11 *hidden neuron* dengan perbandingan data 80 : 20 menunjukkan tingkat akurasi yang tinggi sebesar

40,74%. Teori menjelaskan semakin besar *hidden neuron* yang digunakan, maka akurasi yang diperoleh akan semakin besar (Effendi dkk, 2017).

3.6 Karakterisasi Data

Karakterisasi data dalam penelitian ini bertujuan untuk menentukan konsentrasi *output* model berdasarkan bobot yang sudah teridentifikasi setelah membangun model ANN. Proses karakterisasi data dengan melakukan pengestimasi data berdasarkan data *input* yang digunakan dalam membangun model ANN. Penelitian ini menggunakan tiga data *input* yaitu pH, suhu dan turbiditas untuk mengestimasi konsentrasi parameter *output* yaitu COD dan BOD.

Penetapan nilai *input* disesuaikan dengan rentang nilai yang digunakan dalam membangun model. Persamaan yang digunakan dalam estimasi data sama dengan persamaan yang digunakan dalam fungsi aktivasi yaitu persamaan sigmoid. Perhitungan estimasi bertujuan untuk memperoleh konsentrasi *output* hasil pemodelan berdasarkan bobot hasil pengembangan model. Persamaan sigmoid tersebut sesuai dengan persamaan (11) dan (12)

$$H^* = f(y) = \frac{1}{1 + e^{-y}} \quad (11)$$

$$y = \sum_{i=0}^n X_i W_i \quad (12)$$

