

BAB II

ANALISIS PERMASALAHAN DAN METODOLOGI PENELITIAN

A. Analisis Permasalahan

Prediksi curah hujan adalah permasalahan yang cukup kompleks dan dipengaruhi oleh banyak variabel serta pola musim yang tidak selalu konsisten. Model statistik konvensional seperti regresi linier atau *moving average* memiliki keterbatasan dalam menangkap hubungan non-linear antar variabel.

Kota Makassar dan Kabupaten Maros merupakan wilayah yang sering mengalami perubahan cuaca yang tidak menentu, sehingga sistem prediksi yang akurat dan fleksibel sangat diperlukan untuk memprediksi keadaan tersebut. Salah satu cara yang efektif untuk mengatasi masalah ini yaitu dengan menggunakan *Artificial Neural Network* (ANN).

Namun demikian, keberhasilan ANN tergantung pada proses *network training*, parameter model, dan kualitas data. Oleh karena itu, diperlukan pendekatan yang tepat, salah satunya adalah metode *Backpropagation*, yang merupakan teknik pembelajaran *supervised learning* pada ANN. Permasalahan yang dibahas dalam penelitian ini bukan hanya tentang akurasi prediksi, tetapi juga mengenai bagaimana merancang dan mengimplementasikan ANN secara optimal.

B. Metode Penelitian

Berikut adalah metodologi penelitian yang dilakukan dalam penelitian ini antara lain sebagai berikut.

1. Jenis Penelitian

Jenis penelitian ini adalah penelitian kuantitatif yaitu dilakukan pengumpulan data dan dianalisis untuk mengukur variabel dan hubungannya menggunakan aplikasi perangkat lunak MATLAB untuk membangun model dan mengevaluasi performa model prediksi curah hujan

2. Data dan Sumber Data Penelitian

Data yang digunakan adalah data historis curah hujan bulanan dari tahun 2014 hingga 2024 yang diperoleh dari Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) wilayah Makassar dan Maros.

3. Hipotesis Penelitian

Berdasarkan rumusan masalah dan tujuan penelitian, maka hipotesis yang diajukan dalam penelitian ini adalah sebagai berikut.

a. Umum

Secara umum, diasumsikan bahwa model Jaringan Saraf Tiruan (*Artificial Neural Network*) dengan metode *Backpropagation* mampu memprediksi curah hujan bulanan di wilayah Makassar–Maros dengan tingkat akurasi yang baik. Hal ini didasarkan pada kemampuan ANN dalam mempelajari hubungan nonlinier antara variabel-variabel meteorologis yang kompleks.

b. Khusus

Secara khusus, diasumsikan bahwa terdapat hubungan yang signifikan antara suhu udara, kelembaban udara, penyinaran matahari, dan kecepatan angin terhadap nilai curah hujan bulanan. Model ANN *Backpropagation* diharapkan dapat mengenali pola dan keterkaitan antara keempat parameter tersebut dengan curah hujan, sehingga mampu menghasilkan prediksi yang mendekati data aktual.

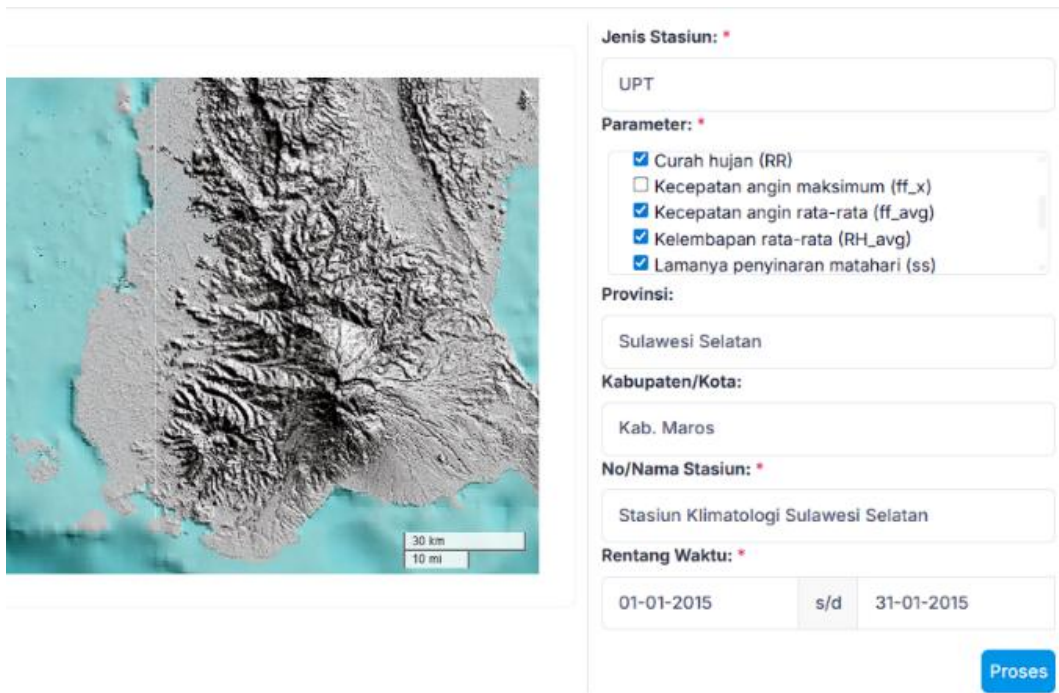
Selain itu, hasil pelatihan jaringan diharapkan menghasilkan nilai *Mean Square Error (MSE)* yang rendah serta koefisien korelasi (*R*) yang tinggi, yang menunjukkan bahwa model memiliki performa prediksi yang baik dan dapat digunakan sebagai pendekatan dalam memprediksi curah hujan.

4. Langkah Penelitian

Langkah-langkah untuk melakukan prediksi pada penelitian ini adalah sebagai berikut:

a. Pengumpulan Data

Data yang diperlukan pada penelitian ini berupa data historis curah hujan harian selama 10 tahun mulai dari 2014 sampai dengan 2024 yang disediakan oleh tiga Stasiun BMKG yaitu Stasiun Meteorologi (BMKG) Paotere, Stasiun Klimatologi Sulawesi Selatan dan Stasiun Meteorologi Klas I Hasanuddin Makassar. Pengumpulan data dilakukan melalui *website* resmi BMKG yang dapat dilihat pada gambar 2.



Sumber: Website BMKG

Gambar 2. Pencarian Data Pada Website BMKG

b. Pra-pemrosesan Data

Setelah dilakukan pengumpulan, data dikategorikan berdasarkan masing-masing data. Variabel curah hujan akan menjadi variabel Y (*output*) dan suhu, kelembaban, penyinaran matahari serta kecepatan angin akan menjadi variabel X (*input*).

Kemudian data curah hujan yang memiliki nilai 0 akan dikeluarkan dan tidak dipakai. Dikeluarkannya nilai nol ini karena nilai nol dapat mengganggu proses pembelajaran jaringan. Pada jaringan saraf tiruan, setiap data input akan dikalikan dengan bobot (*weight*) dan diolah melalui fungsi aktivasi seperti sigmoid. Jika terlalu banyak data yang bernilai nol (0), maka hasil perkalian awal dari neuron juga akan menjadi nol, sehingga sinyal yang diteruskan ke lapisan berikutnya menjadi sangat kecil atau

bahkan hilang (*vanishing gradient*), bobot tidak mengalami pembaruan secara efektif saat proses *backpropagation* dan jaringan kesulitan belajar pola yang sebenarnya dari data.

Tabel 1. Format Data Curah Hujan Bulanan Tahun 2014 Hingga 2024

<u>Tahun</u>	Bulan	<u>Curah Hujan</u>	<u>Suhu</u>	<u>kelembaban</u>	<u>Penyinaran Matahari</u>	<u>Kecepatan Angin</u>
		(mm)	(°C)	(%)	(Jam)	(m/s)
2014	Januari	23.87	24.98	88.74	2.42	3.29
	<u>Februari</u>	10.99	24.62	78.03	3.83	2.23
	Maret	10.02	27.63	85.65	4.85	2.10
	April	8.97	27.18	81.68	5.54	2.06
	Mei	3.26	24.03	68.00	6.65	1.87
	Juni	4.30	21.84	63.61	5.70	1.77
	Juli	0.96	16.03	44.87	6.56	1.94
	Augustus	0.17	25.63	70.32	7.45	2.10
	<u>september</u>	0.00	24.12	59.81	8.39	2.03
	<u>Oktober</u>	0.00	11.37	26.29	5.40	1.81
	November	2.98	22.40	60.61	7.45	2.23
	<u>Desember</u>	19.14	14.89	48.03	4.07	2.52

c. Mengeluarkan Data *Outlier*

Dalam data-data yang sudah dikumpulkan ditemukan juga beberapa data yang memiliki nilai terlampaui besar sehingga bisa saja mempengaruhi proses *network training* dan nilai keluaran nantinya. Sihombing (2022) menjelaskan data *outlier* merupakan data-data yang bobotnya bisa didefinisikan sebagai bobot menyimpang terlalu jauh dari data-data yang lain. Hal ini nantinya dapat menyebabkan data yang ada tidak terdistribusi secara normal.

Oleh karena itu digunakan salah satu metode untuk mengeluarkan data *outlier* ini yaitu metode *Z-score* yang merupakan ukuran statistik yang

digunakan untuk menentukan seberapa jauh suatu nilai data menyimpang dari rata-ratanya dalam satuan simpangan baku. Secara matematis, rumus *Z-score* cukup sederhana:

$$Z_i = \frac{X_i - \bar{X}}{s} \quad (1)$$

Dengan keterangan:

Z_i : nilai *Z-score* dari data ke- i ,

X_i : nilai data ke- i ,

$\bar{X}$: rata-rata (*mean*) dari keseluruhan data,

s : simpangan baku (*standard deviation*) dari data.

Rentang *Z-score* yang digunakan adalah -3 sampai dengan 3 yang didasarkan pada prinsip distribusi normal (Gaussian Distribution), dimana sekitar 99,7% dari seluruh data akan berada dalam rentang tersebut. Ghozali (2018) mengatakan untuk sampel yang lebih besar (lebih dari 80), standar skor dinyatakan outlier apabila nilainya berada pada kisaran 3 ke atas.

d. Normalisasi Data

Normalisasi data ini dilakukan untuk menyamakan ukuran nilai dari tiap data agar bobot datanya menjadi lebih seimbang. Dalam penelitian ini digunakan teknik normalisasi *min-max* berjarak antara 0 sampai 1. Digunakan *range* 0 sampai 1 karena menurut Chamidah (2012), fungsi sigmoid yang biasanya digunakan dalam *backpropagation* akan membawa nilai *input* dengan range yang tak terbatas ke nilai *output* yang terbatas, yaitu dalam range 0 sampai 1. Supaya dapat membawa range nilai *output* ke dalam range *input*, maka pada data *input* harus dilakukan normalisasi

data ke dalam range 0 sampai 1, sehingga *outputnya* dapat di denormalisasi ke dalam range nilai *input*. Menurut Han, et al. (2011), normalisasi data teknik *min-max* merupakan transformasi linier dari data asli. Persamaan normalisasi data *min-max* dapat dilihat pada persamaan 1.

$$v'_i = \frac{v_i - \min_A}{\max_A - \min_A} (\text{new_max}_A - \text{new_min}_A) + \text{new_min}_A \quad (2)$$

Lalu untuk melakukan denormalisasi nantinya, dapat dilihat pada persamaan 2 berikut.

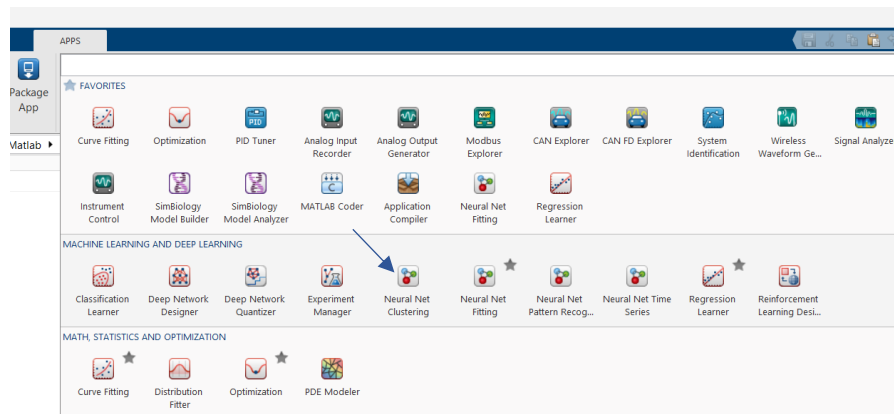
$$v_i = \frac{v'_i - \text{new_min}_A}{(\text{new_max}_A - \text{new_min}_A)} (\max_A - \min_A) + \min_A \quad (3)$$

Di mana :

- v : nilai normal
- A : variabel
- $\min_A$: nilai minimum pada variabel
- $\max_A$: nilai maksimum pada variabel
- new_min_A : nilai minimum baru pada nilai normal
- new_max_A : nilai maksimum baru pada nilai normal

e. Desain, *Network training* dan Pengujian model ANN

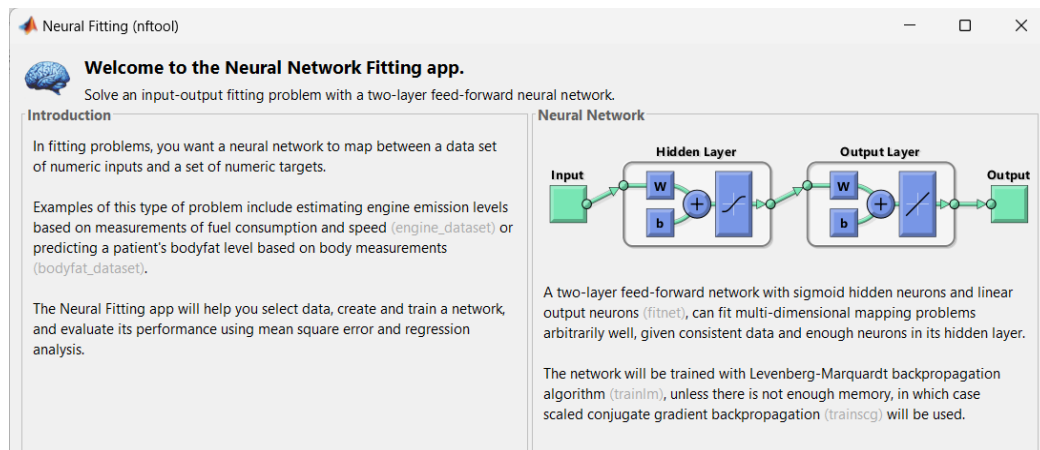
Pada Langkah ini, dilakukan pemodelan arsitektur ANN dengan menggunakan MATLAB. Buka aplikasi MATLAB, kemudian pilih *apps* dan *pilih neural net fitting*.



Sumber: MATLAB R2021a

Gambar 3. Tampilan Menu Apps Pada MATLAB

setelah itu, akan muncul tampilan dari program *neural net fitting* seperti pada gambar 4.



Sumber: MATLAB R2021a

Gambar 4. App Neural Net Fitting

f. Validasi Kinerja

Perlu dilakukan validasi kinerja pada jaringan yang telah dijalankan karena dalam pengukuran prediksi, tidak ada yang dapat menyentuh hingga

100%. Oleh karena itu, diperlukan suatu metode perhitungan untuk mengetahui tingkat kesalahan yang terjadi.

(i) Rata-rata Kuadrat Kesalahan (*Mean Square Error* = MSE)

Voni (2016), menyebutkan bahwa *mean square error* dihitung dari penjumlahan kuadrat semua kesalahan prediksi pada setiap periode dan membaginya dengan jumlah periode peramalan. Rumus dari MSE ini dapat dilihat pada persamaan 3.

$$MSE = \frac{(A_t - F_t)^2}{n} \quad (4)$$

Di mana :

A_t = Nilai curah hujan aktual

F_t = Nilai hasil prediksi curah hujan

n = jumlah data

(ii) Rata-rata presentase kesalahan absolut (*Mean Absolute Presentage Error* = MAPE)

Voni (2016) juga menyebutkan MAPE merupakan perhitungan yang membandingkan presentase perbedaan nilai rata-rata absolut antara nilai yang diprediksi dengan nilai aktual yang persamaannya dapat dilihat pada persamaan 4 berikut.

$$MAPE = \left(\frac{1}{n} \right) \left| \frac{F_t - A_t}{A_t} \right| \quad (5)$$

Di mana :

A_t = Nilai curah hujan aktual

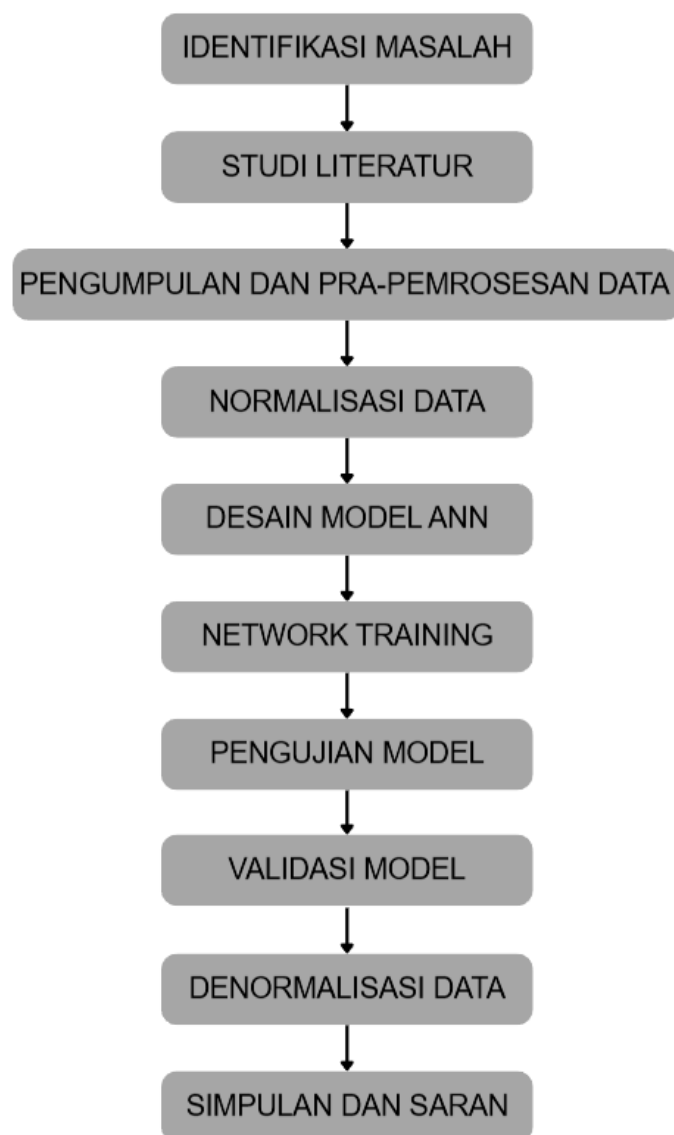
F_t = Nilai hasil prediksi curah hujan

n = jumlah data

Jadi, secara garis besar Validasi kinerja model jaringan saraf tiruan dilakukan menggunakan indikator *Mean Squared Error* (MSE) dan *Mean Absolute Percentage Error* (MAPE).

C. Alur Penelitian

Alur penelitian ini terdiri dari beberapa tahapan yang dijelaskan dalam diagram alur berikut:



Gambar 5. Alur Penelitian