

SELEKSI FITUR MENGGUNAKAN BACKWARD ELIMINATION PADA PREDIKSI CUACA DENGAN NEURAL NETWORK

Yunita

AMIK BSI Jakarta

yunita.ynt@bsi.ac.id

Abstrak

Prakiraan Cuaca merupakan bagian yang penting untuk pengambilan keputusan seperti kontrol lalu lintas udara, kapal routing, pertanian, pembangkit listrik dan keuangan. Oleh karena itu banyak pihak yang membutuhkan informasi kondisi atmosfer (cuaca) yang lebih cepat, lengkap, dan akurat. Metode Neural Network lebih efisien dalam perhitungan cepat dan mampu menangani data yang tidak stabil yang khas dalam kasus data prakiraan cuaca, tetapi Neural Network menghadapi keterbatasan dalam mempelajari pola klasifikasi jika dataset memiliki data yang luar biasa dan dimensi yang kompleks. Untuk menyelesaikan masalah tersebut diperlukan metode penyeleksian fitur. Metode yang digunakan adalah Backward Elimination Untuk Seleksi Fitur Pada Metode Neural Network. Untuk Prediksi Cuaca dengan input data adalah data sinoptik. Beberapa percobaan dilakukan untuk mendapatkan arsitektur yang optimal dan menghasilkan prediksi yang akurat. Hasil penelitian menunjukkan metode jaringan syaraf tiruan berbasis backward elimination menghasilkan peningkatan akurasi 4% dibandingkan hanya dengan menggunakan metode jaringan syaraf tiruan saja. Melalui penelitian ini dapat diketahui sejauh mana metode Neural Network menggunakan seleksi fitur Backward Elimination mempunyai tingkat akurasi lebih baik sehingga dapat membantu banyak pihak yang membutuhkan informasi kondisi atmosfer (cuaca) yang lebih cepat, lengkap, dan akurat.

Kata Kunci: prediksi, cuaca, Jaringan Syaraf Tiruan, Backward Elimination.

Abstract

Weather Forecast is an important part of the decision making such as air traffic control, ship routing, agriculture, power generation and financial. Therefore, many people who need information atmospheric conditions (weather) is more rapid, complete, and accurate. Neural Network method is more efficient in calculations quickly and is able to handle data that is not stable in the case of typical weather forecast data, but Neural Network face limitations in studying the pattern classification if the dataset has exceptional data and complex dimensions. To resolve these problems required method of selecting features. The method used is the Backward Elimination for Seleksi Fitur Method of Neural Network On. For weather prediction with the data input is data synoptic. Several experiments were conducted to obtain the optimal architecture and generate accurate predictions. The results show the method of artificial Neural Network-based backward elimination resulted in increased accuracy compared to only 4% using Neural Network. Through this research can be known to what extent the method of Neural Network using Backward Elimination feature selection has a better accuracy so that it can help many parties who require information of the condition of the atmosphere (weather) faster, complete, and accurate.

Keywords: Prediction, Weather, Neural Network, Backward Elimination.

1. PENDAHULUAN

Menurut Fraley, Raftery, Gneiting, Sloughter & Berrocal (2011) Selama dua dekade terakhir, prakiraan cuaca mengalami pergeseran paradigma terhadap probabilistik perkiraan, sehingga diperlukan prakiraan cuaca dengan

probabilitik yang optimal. Cuaca merupakan bagian yang penting dalam mengambil keputusan seperti kontrol lalu lintas udara, kapal routing, pertanian, pembangkit listrik dan keuangan. Oleh karena itu banyak pihak yang membutuhkan informasi kondisi atmosfer

(cuaca) yang lebih cepat, lengkap, dan akurat.

Penelitian terdahulu mengenai prediksi cuaca dalam suatu wilayah telah diselesaikan dengan beberapa metode seperti Numerical weather Prediction (NWP) (Simeonov et al., 2007), Adaptive Neuro Fuzzy Inference System (ANFIS) (Aldrian & Djamil, 2012). Neural Network (NN) (Hung, Babel, Weesakul & Tripathi, 2009) untuk prediksi cuaca. Numerical Weather Prediction (NWP) memiliki kelebihan memberikan prediksi pada banyak variabel atmosfer seperti suhu, tekanan, angin, dan curah hujan dengan penggunaan program komputer yang kompleks tetapi NWP memiliki kelemahan mensimulasi model dengan data yang ada sehingga menyebabkan ketidak akuratan dalam memprediksi (Simeonov, Kilifarev, & Ilarionov, 2007). Penelitian berikutnya yang dilakukan oleh (Aldrian & Djamil, 2008) menggunakan metode ANFIS melakukan peramalan cuaca jangka pendek menggunakan data tujuh bulan (januari - Juli 2005) di Timika, Studi ini menunjukkan bahwa model ANFIS memiliki keunggulan yakni memiliki sensitifitas yang baik terhadap besaran yang berbeda dan ukuran skala, tetapi ANFIS bukan alat yang tepat untuk proses stokastik seperti prakiraan cuaca.

Penelitian yang dilakukan oleh (Hung et al., 2009) menggunakan Neural Network (NN) dengan menggunakan data dari tahun 1991 sampai dengan tahun 2003 yang diambil dari 51 pos hujan di Bangkok, studi ini menunjukkan bahwa NN lebih efisien dalam perhitungan cepat dan mampu menangani data yang tidak stabil yang khas dalam kasus data untuk prakiraan cuaca. tetapi NN memiliki kekurangan jika digunakan untuk peramalan dalam jangka panjang. Neural Network menurut (Lin, Chen, Wu, & Chen, 2009) menghadapi keterbatasan dalam mempelajari pola klasifikasi Jika dataset memiliki data yang luar biasa dan dimensi tinggi. Untuk menyelesaikan permasalahan dalam keterbatasan dikarenakan dataset yang besar dan dimensi yang tinggi maka diperlukan sebuah metode penyeleksian fitur agar Neural Network dapat menghasilkan prediksi yang akurat. Seleksi fitur adalah masalah yang berkaitan erat dengan pengurangan dimensi. Tujuan dari seleksi adalah untuk mengidentifikasi fitur dalam kumpulan data yang sama pentingnya,

dan membuang semua fitur lain seperti informasi yang tidak relevan dan berlebihan. sehingga memungkinkan operasi yang lebih cepat dan lebih efektif (Beniwal & Arora, 2012). Seleksi Fitur terdiri dari dua buah metode yaitu metode wrapper dan metode filter. Metode wrapper melakukan seleksi fitur dengan menggunakan kinerja dari algoritma learning, sedangkan metode filter melakukan proses seleksi fitur terlebih dahulu kemudian melakukan algoritma pembelajarannya (Beniwal & Arora, 2012). Metode wrapper langsung menggunakan algoritma induksi untuk mengevaluasi subset fitur sehingga lebih unggul dalam proses prediksi (Zhu, Ong & Dash, 2007). Metode wrapper terdiri dari dua yaitu Forward Selection dan Backward elimination, backward elimination memiliki kelebihan dalam hal komputasi data dibandingkan metode forward selection (Kohavi & John, 1997). Pada penelitian ini Seleksi Fitur dengan metode wrapper menggunakan algoritma Backward Elimination akan diterapkan untuk pemilihan subset fitur yang sesuai pada Neural Network, sehingga hasil prediksi lebih akurat.

Prediksi Cuaca

Pengertian Prediksi Cuaca Cuaca merupakan bentuk awal yang dihubungkan dengan penafsiran dan pengertian akan kondisi fisik udara sesaat pada suatu lokasi dan suatu waktu, sedangkan iklim merupakan kondisi lanjutan dan merupakan kumpulan dari kondisi cuaca yang kemudian disusun dan dihitung dalam bentuk rata-rata kondisi cuaca dalam kurun waktu tertentu (Winarso, 2003). Cuaca terbentuk dari gabungan. Unsur cuaca dan jangka waktu cuaca bias hanya beberapa jam saja.

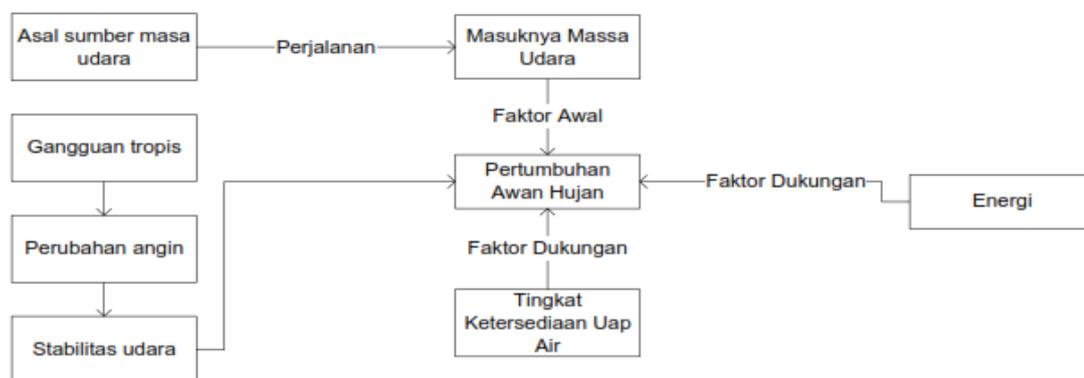
Misalnya pagi hari, siang hari, sore hari atau malam hari dan keadaannya bisa berbeda-beda untuk setiap tempat serta setiap jamnya. Di Indonesia keadaan cuaca selalu diumumkan untuk jangka waktu sekitar 24 jam melalui prediksi cuaca yang dikembangkan oleh Badan Meteorologi Klimatologi dan Geofisika (BMKG), Departemen Perhubungan. Ilmu cuaca atau meteorology adalah ilmu pengetahuan yang mengkaji peristiwa-peristiwa cuaca dalam jangka waktu dan ruang terbatas. Sedangkan ilmu iklim atau klimatologi adalah ilmu pengetahuan yang juga

mengkaji tentang gejala-gejala cuaca tetapi sifat-sifat dan gejala-gejala tersebut mempunyai sifat umum dalam jangka waktu dan daerah yang luas di atmosfer permukaan bumi (Winarso, 2003). Prediksi cuaca merupakan rangkuman informasi kondisi cuaca harian hingga mingguan, sedangkan prediksi iklim umumnya merupakan unsur – unsur iklim yang umumnya untuk wilayah Indonesia adalah prediksi hujan bulanan atau prediksi hujan yang berlangsung dalam satu musim (Georgious et al, 2008). Di Indonesia informasi prakiraan cuaca yang sudah dikenal oleh masyarakat adalah berawan, cerah dan hujan. Sementara itu terjadinya hujan dikaitkan dengan proses fisis dan dinamika atmosfer yang diketahui melalui parameter-parameternya seperti adanya massa udara, gaya vertical dan

energy. Keadaan cuaca di Indonesia juga dipengaruhi oleh : a. Fenomena badai tropis atau vortex. Keberadaan siklon tropis akan mengganggu sistem cuaca dari keadaan semula b. Palung, dapat diketahui melalui citra satelit yang ditandai dengan sederetan awan yang memanjang. c. Konvergensi (daerah pertemuan angin dimana kecepatan angin semakin kecil) d. Shearline (daerah belokan angin dimana kecepatan anginnya lebih rendah dibandingkan dengan daerah sekitarnya).

Prinsip Dasar Prediksi Cuaca

Secara umum prinsip dasar prediksi cuaca yang sudah lama dilakukan oleh prakirawan BMKG adalah seperti gambar berikut :



Gambar 1. Prinsip Prakiraan Cuaca

Seperti digambarkan pada Gambar 1. Prinsip Prakiraan cuaca di Indonesia dipengaruhi oleh jenis masa udara yang terdapat disekitar wilayah Indonesia karena itu perlu mengetahui pola tekanan udara yang ada di Asia dan Australia, Masuknya massa udara menjadi faktor awal pertumbuhan awan hujan, selain faktor awal gangguan tropis, perubahan angin, stabilitas udara juga menjadi faktor pertumbuhan awan hujan. Selain faktor awal ada juga faktor dukungan yang diperoleh dari tingkat ketersediaan uap air dan energi. Metode yang saat ini dipakai pada Badan Meteorologi Klimatologi dan Geofisika (BMKG) Pondok Betung adalah metode dengan menginterpretasikan data hasil pengamatan dan data model prakiraan cuaca atau NWP (Numerical Weather Prediction). Metode ini mempunyai kelemahan, karena setiap prakirawan akan menghasilkan interpretasi yang berbeda beda, dengan demikian metode ini menjadi subjektif.

Algoritma Neural Network

Menurut (Graupe, 2007) dalam buku Heaton Neural Network adalah jaringan saraf yang mensimulasikan jaringan saraf biologis manusia kedalam arsitektur komputer dan arsitektur algoritma baru terhadap komputer konvensional. Hal ini memungkinkan penggunaan operasi komputasi (penambahan, pengurangan, dan elemen logika fundamental) yang sangat sederhana untuk memecahkan masalah yang kompleks, matematis yang tidak jelas, masalah nonlinear atau masalah stokastik.

Algoritma Backpropagation

Pembelajaran

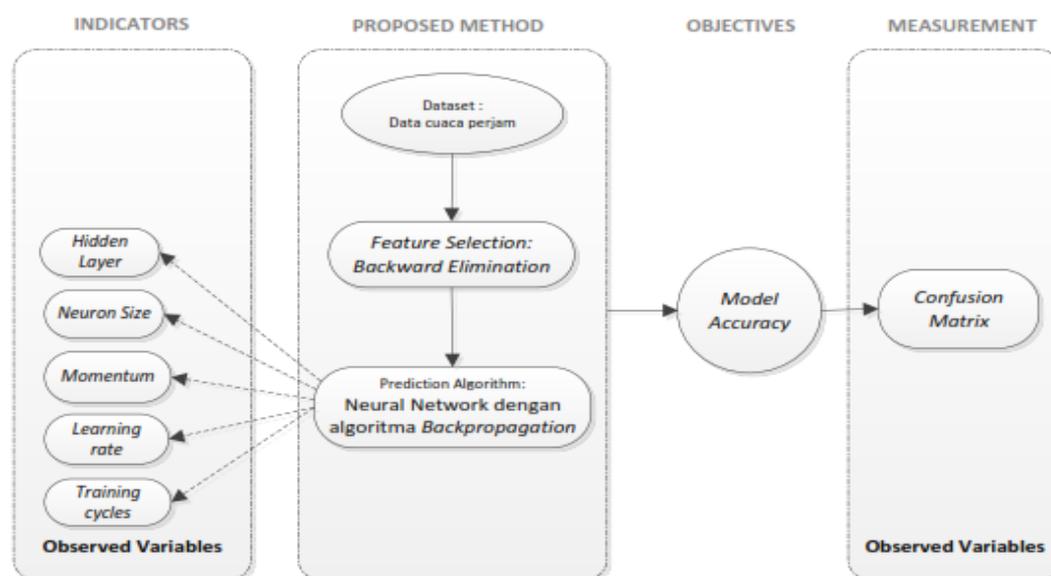
Backpropagation merupakan algoritma pembelajaran yang terawasi dan backpropagation merupakan metode pembelajaran jaringan syaraf tiruan yang paling umum digunakan dan bekerja melalui proses secara iteratif dengan menggunakan sekumpulan contoh data

(data training), membandingkan nilai prediksi dari jaringan dengan setiap contoh data. Dalam setiap proses, bobot relasi dalam jaringan dimodifikasi untuk meminimalkan nilai Mean Square Error (MSE) antara nilai prediksi dari jaringan syaraf tersebut dilakukan dalam arah mundur, dari output layer hingga layer pertama dari hidden layer sehingga metode ini disebut backpropagation (Jong Jek Siang, 2009).

Seleksi Fitur

Seleksi Fitur adalah masalah berkaitan erat dengan pengurangan dimensi. Tujuan Seleksi Fitur adalah untuk mengidentifikasi fitur dalam kumpulan data yang sama pentingnya, dan membuang semua fitur lain seperti informasi yang tidak relevan dan berlebihan. Karena Seleksi Fitur mengurangi dimensi dari data, sehingga memungkinkan operasi yang lebih efektif & algoritma data mining yang cepat (yaitu data algoritma Mining dapat dioperasikan lebih cepat dan lebih efektif dengan menggunakan Seleksi Fitur) (Maimon & Rokach, 2010). Tujuan dari seleksi fitur adalah untuk mengidentifikasi dan mencari variable yang paling relevan dan variable yang tidak redundan dari subset variable yang potensial (Mao, 2004). Tujuan seleksi fitur adalah untuk pengurangan fitur, untuk menghilangkan dari dataset subset dari variabel yang tidak dianggap relevan (Vercellis, 2009).

Kerangka Pemikiran



Gambar 2. Kerangka Pemikiran

Algoritma Backward Elimination

Backward Elimination konvensional menurut adalah menghilangkan sebuah model regressor satu persatu berdasarkan penurunan dalam sebuah model yang fit (Hong, Harris, Brown, & Chen, 2002).

Berikut langkah-langkah dalam backward elimination :

- Mulailah dengan semua prediktor dalam model
- Hapus prediktor dengan p-value tertinggi atau lebih besar dari target
- Perbaiki model dan ulangi langkah ke-2.
- Hentikan ketika semua jika p-value kurang dari target.

Metode Evaluasi Performa Prediksi Pengujian K-Fold Cross

Validation Cross Validation adalah teknik validasi dengan membagi data secara acak kedalam k bagian dan masing-masing bagian akan dilakukan proses klasifikasi (Han & Kamber, 2006). Dengan menggunakan cross validation akan dilakukan percobaan sebanyak k. Data yang digunakan dalam percobaan ini adalah data training untuk mencari nilai error rate secara keseluruhan. Secara umum pengujian nilai k dilakukan sebanyak 10 kali untuk memperkirakan akurasi estimasi. Dalam penelitian ini nilai k yang digunakan berjumlah 10 atau 10-fold cross validation.

Dataset yang digunakan dalam penelitian ini menggunakan data cuaca perjam dari tahun 2010 s/d tahun 2013 yang direkam dari radio sonde yang didapat dari stasiun klimatologi. Pada tahap preprocessing menggunakan algoritma Backward Elimination (Proposed Methode) dengan memproses variable input dan target dataset. Setiap dihasilkan sebuah bagian maka akan diseleksi dengan algoritma Neural Network dengan hidden layer, neuron size, momentum, learning rate dan training cycle sebagai indicators. Pengukuran (Measurement) yang digunakan untuk menilai keakuratan dan optimal atau tidaknya hasil prediksi ini, ditentukan oleh nilai confusion matrix (CM). Dimana, semakin tinggi nilai dari CM maka semakin akurat prediksi yang dihasilkan.

2. METODE PENELITIAN

Perancangan Penelitian

Penelitian ini dilakukan dengan tujuan untuk menyediakan informasi cuaca yang akurat sehingga dapat memberikan informasi sedini mungkin kepada pihak yang membutuhkan dan untuk menangani permasalahan banjir. Metode penelitian ini adalah penelitian eksperimen dengan tahapan penelitian sebagai berikut : 1. Pengumpulan Data Pada bagian ini dijelaskan tentang bagaimana dan darimana data dalam penelitian ini didapatkan, meliputi data sekunder dan data primer. 2. Pengolahan Awal Data Pada bagian ini dijelaskan tentang tahap awal data mining. Pengolahan awal data meliputi proses input data ke format yang dibutuhkan, pengelompokan dan penentuan atribut data. 3. Metode yang Diusulkan Pada bagian ini dijelaskan tentang metode yang diusulkan untuk digunakan pada prediksi cuaca jangka pendek. Penjelasan meliputi pengaturan dan pemilihan nilai dari parameter-parameter dan arsitektur melalui uji coba. 4. Eksperimen dan Pengujian Metode Pada bagian ini dijelaskan tentang langkah-langkah eksperimen meliputi cara pemilihan arsitektur yang tepat dari model atau metode yang diusulkan sehingga didapatkan hasil yang dapat membuktikan bahwa metode yang digunakan adalah

tepat. 5. Evaluasi dan Validasi Hasil Evaluasi dilakukan dengan mengamati hasil prediksi menggunakan Algoritma Soft Computing. Validasi dilakukan dengan membandingkan nilai error hasil prediksi masing-masing algoritma sehingga dapat diketahui algoritma yang lebih akurat.

Metode Pengumpulan Data

Metode Pengumpulan data yang digunakan dalam penelitian ini terdiri dari data:

- a. Data primer yaitu data yang dihasilkan dari pengolahan data dengan menggunakan Rapid Miner.
- b. Data sekunder berisi data atribut penunjang dalam prediksi cuaca yang didapat dari stasiun klimatologi BMKG.

Data yang digunakan dalam penelitian ini merupakan data yang diambil dari Stasiun klimatologi BMKG. Observasi dilakukan pada stasiun tersebut untuk mengumpulkan data cuaca. Data yang dikumpulkan diambil dari sejumlah data sinoptik dan data yang mendukung untuk prediksi cuaca, terdapat 14 buah atribut, ke-14 data tersebut terupdate setiap 1 jam sekali selama 12 jam dimulai dari jam (07.00 s/d 19:00). Data diambil pada bulan Januari 2010.

Data penelitian diperoleh dari stasiun klimatologi BMKG berupa data synoptik sebanyak 22 attribut, tetapi peneliti hanya menggunakan sebanyak 14 attribute karena attribut tersebut tidak relevan untuk penelitian ini.

Atribut yang digunakan untuk pengklasifikasian dalam prediksi cuaca terdapat 14 buah atribut yang terdiri dari : Titik embun, jumlah awan, arah angin, kecepatan angin, penglihatan datar, tekanan udara, suhu, jumlah awan rendah, jenis awan, tinggi dasar awan, awan tengah, kelembaban, tekanan udara stasiun, temperatur pada saat pengawasan.

3. PEMBAHASAN

a. Hasil Eksperimen

3.1.1. Pengolahan Data

Langkah pertama yang dilakukan adalah mengolah data yang ada, data awal yang didapat dari stasiun klimatologi adalah seperti Tabel 1.

Tabel 1. Data awal yang belum diolah

T _d T _d T _d	N	Dd	ff	VV	QFF	T _i T _i T _i	N _h	C _L	h	C _M	UU	QFE	T _w T _w T _w	ww
235	0	00	00	00	0093	252	0	0	0	0	90	0063	240	cerah
236	0	00	00	00	0094	270	0	0	0	0	83	0064	248	cerah
232	0	00	00	00	0094	286	0	0	0	0	80	0064	250	cerah
242	0	00	00	00	0095	300	0	0	0	0	72	0065	260	cerah
230	0	00	00	00	0095	310	0	0	0	0	66	0065	260	cerah
236	0	00	00	00	0092	324	0	0	0	0	64	0062	260	cerah
234	0	00	00	00	0080	324	0	0	0	0	60	0050	262	cerah
238	7	27	02	60	0067	322	7	0	9	3	62	0037	264	berawan
256	7	31	02	60	0061	326	5	2	5	3	67	0031	276	berawan
254	7	27	06	60	0058	322	1	1	5	3	68	0028	274	hujan
227	8	27	04	58	0063	310	2	5	5	7	62	0033	254	hujan

Berikut perhitungan transform dengan menggunakan fungsi aktivasi sigmoid: Berdasarkan data sinoptik cuaca harian diketahui nilai minimum (a) = 214 nilai maximum (b) = 266 data yang akan di transform (x) = 235.

$$x^1 = \frac{0.8(x - a)}{b - a} + 0.1$$

$$x^1 = \frac{0.8(235 - 214)}{(266 - 214)} + 0.1$$

$$x^1 = 0.4$$

Tabel 2. Data Training setelah ditransform

T _d T _d T _d	N	dd	Ff	VV	QFF	T _i T _i T _i	N _h	C _L	h	C _M	UU	QFE	T _w T _w T _w	ww
0.4	0.1	0.1	0.1	0.1	0.1	0.3	0.1	0.1	0.1	0.1	0.8	0.5	0.4	cerah
0.4	0.1	0.1	0.1	0.1	0.1	0.4	0.1	0.1	0.1	0.1	0.7	0.5	0.5	cerah
0.4	0.1	0.1	0.1	0.1	0.1	0.5	0.1	0.1	0.1	0.1	0.6	0.5	0.5	cerah
0.5	0.1	0.1	0.1	0.1	0.1	0.6	0.1	0.1	0.1	0.1	0.5	0.5	0.7	cerah
0.3	0.1	0.1	0.1	0.1	0.1	0.7	0.1	0.1	0.1	0.1	0.3	0.5	0.7	cerah
0.4	0.1	0.1	0.1	0.1	0.1	0.8	0.1	0.1	0.1	0.1	0.3	0.4	0.7	cerah
0.4	0.1	0.1	0.1	0.1	0.1	0.8	0.1	0.1	0.1	0.1	0.2	0.3	0.7	cerah
0.5	0.8	0.7	0.7	0.3	0.9	0.8	0.8	0.1	0.9	0.4	0.3	0.2	0.7	berawan
0.7	0.8	0.8	0.8	0.3	0.9	0.8	0.6	0.3	0.5	0.4	0.4	0.1	0.9	berawan
0.7	0.8	0.7	0.7	0.6	0.9	0.8	0.2	0.2	0.5	0.4	0.4	0.1	0.8	hujan
0.3	0.9	0.7	0.7	0.5	0.9	0.7	0.3	0.7	0.5	0.9	0.3	0.2	0.6	hujan
0.3	0.7	0.6	0.6	0.3	0.8	0.5	0.3	0.7	0.5	0.9	0.4	0.2	0.5	berawan
0.4	0.7	0.6	0.6	0.3	0.6	0.5	0.3	0.7	0.5	0.9	0.5	0.3	0.5	berawan
0.4	0.9	0.1	0.1	0.1	0.9	0.3	0.4	0.7	0.5	0.9	0.7	0.5	0.4	berawan
0.4	0.8	0.7	0.7	0.3	0.9	0.5	0.8	0.1	0.9	0.4	0.5	0.5	0.5	berawan
0.4	0.7	0.7	0.7	0.3	0.9	0.7	0.4	0.1	0.9	0.4	0.3	0.6	0.6	berawan
0.4	0.5	0.7	0.7	0.3	0.9	0.8	0.2	0.2	0.5	0.1	0.3	0.6	0.7	cerah
0.4	0.5	0.6	0.6	0.3	0.9	0.8	0.3	0.2	0.5	0.1	0.2	0.5	0.7	cerah
0.4	0.5	0.7	0.7	0.6	0.9	0.9	0.3	0.2	0.5	0.1	0.2	0.4	0.8	cerah
0.4	0.7	0.7	0.7	0.5	0.9	0.9	0.5	0.2	0.5	0.1	0.1	0.4	0.8	berawan

Kemudian data tersebut dimasukan data menjadi seperti yang terlihat pada kedalam framework RapidMiner maka Tabel 3.

Tabel 3. Data Training yang dipakai pada framework RapidMiner

Role	Name	Type	Statistics	Range	Missings
label	O	polynomial	mode = berawan (169), cerah (115), berawan (1		0
regular	A	numeric	avg = 0.527 +/- 0.127	[0.100 ; 1.146]	0
regular	B	numeric	avg = 0.733 +/- 0.161	[0.100 ; 0.900]	0
regular	C	numeric	avg = 1.110 +/- 11.050	[0.100 ; 222.300]	0
regular	D	numeric	avg = 1.110 +/- 11.050	[0.100 ; 222.300]	0
regular	E	numeric	avg = 0.326 +/- 0.205	[0.100 ; 0.989]	0
regular	F	numeric	avg = 0.852 +/- 0.127	[0.100 ; 0.900]	0
regular	G	numeric	avg = 0.528 +/- 0.166	[0.100 ; 0.900]	0
regular	H	numeric	avg = 0.527 +/- 0.193	[0.100 ; 0.900]	0
regular	I	numeric	avg = 0.449 +/- 0.245	[0.100 ; 0.900]	0
regular	J	numeric	avg = 0.574 +/- 0.176	[0.100 ; 0.900]	0
regular	K	numeric	avg = 0.351 +/- 0.180	[0.100 ; 1.014]	0
regular	L	numeric	avg = 0.567 +/- 0.193	[0.100 ; 0.937]	0
regular	M	numeric	avg = 0.609 +/- 0.189	[0.100 ; 1.013]	0
regular	N	numeric	avg = 0.610 +/- 0.138	[0.100 ; 0.952]	0

3.1.2. Neural Network

3.1.2.1 Parameter Neural Network

Eksperimen pertama dilakukan untuk mendapatkan nilai training cycles, yaitu dilakukan uji coba dengan memasukan nilai dengan range 100 sampai dengan 1000, nilai 0.1 untuk learning rate dan nilai 0.1 untuk momentum. Berikut hasil eksperimen terhadap training cycles.

Tabel 4. eksperimen untuk menentukan nilai training cycles

Training Cycle	Learning Rate	Momen tum	Accuracy
100	0.1	0.1	71.99%
200	0.1	0.1	71.49%
300	0.1	0.1	71.25%
400	0.1	0.1	72.24%
500	0.1	0.1	72.26%
600	0.1	0.1	71.50%
700	0.1	0.1	70.74%
800	0.1	0.1	69.75%
900	0.1	0.1	69.01%
1000	0.1	0.1	69.76%

Hasil penelitian menunjukkan bahwa menambah nilai training cycles hanya membawa perubahan sedikit terhadap nilai root mean square error (RMSE), berdasarkan dari hasil training cycle diatas training cycle dengan nilai 500 memiliki nilai akurasi tertinggi.

Tabel 5. eksperimen untuk menentukan nilai learning rate

Training Cycle	Learnin g Rate	Momen tum	Accuracy
500	0.1	0.1	72.26%
500	0.2	0.1	71.23%

500	0.3	0.1	69.24%
500	0.4	0.1	69.50%
500	0.5	0.1	69.26%

Nilai learning rate yang dipilih pada hasil eksperimen diatas adalah nilai learning rate dengan nilai akurasi tertinggi yaitu dipilih nilai 0.1 untuk learning rate di eksperimen selanjutnya yaitu menentukan nilai momentum.

Tabel 6. eksperimen untuk menentukan nilai momentum

Training Cycle	Learning Rate	Momen tum	Accuracy
500	0.1	0.0	71.26%
500	0.1	0.1	72.26%
500	0.1	0.2	71.74%
500	0.1	0.3	69.75%
500	0.1	0.4	69.76%
500	0.1	0.5	69.74%
500	0.1	0.6	71.23%
500	0.1	0.7	70.75%
500	0.1	0.8	69.76%
500	0.1	0.9	70.26%

Dari hasil eksperimen untuk menentukan nilai momentum dengan nilai akurasi tertinggi didapat pada saat momentum bernilai 0.1.

3.1.2.2. Menentukan Jumlah Hidden Layer dan Hidden Layer Sizes

Pada jumlah hidden layer sebanyak 1, dilakukan percobaan pada hidden layer sizes dengan range 1 sampai dengan 10 sizes. Berikut hasil eksperimen dengan 1 hidden layer.

Tabel 7. Eksperimen untuk menentukan Hidden Layer Sizes pada Hidden Layer 1

Training Cycle	Learning Rate	Momentum	Sizes	Accuracy
100	0.1	0.1	1	62.03%
200	0.1	0.1	2	71.72%
300	0.1	0.1	3	70.98%
400	0.1	0.1	4	71.47%
500	0.1	0.1	5	71.71%
600	0.1	0.1	6	68.78%
700	0.1	0.1	7	71.98%
800	0.1	0.1	8	70.49%
900	0.1	0.1	9	72.97%
1000	0.1	0.1	10	72.26%

Hasil eksperimen menunjukkan 1 hidden layer dengan sizes 9 yang menghasilkan nilai akurasi tertinggi yaitu, dengan nilai 72.97 %.

Tabel 8. eksperimen untuk menentukan Hidden Layer Sizes pada Hidden Layer 2

Training Cycles	Learning rate	Momentum	Size h1	Sizes h2	Accuracy
500	0.1	0.1	9	1	61.28%
500	0.1	0.1	9	2	66.52%
500	0.1	0.1	9	3	69.51%
500	0.1	0.1	9	4	70.98%
500	0.1	0.1	9	5	69.26%
500	0.1	0.1	9	6	68.26%
500	0.1	0.1	9	7	70.22%
500	0.1	0.1	9	8	70.24%
500	0.1	0.1	9	9	70%
500	0.1	0.1	9	10	69.49%

Hasil eksperimen menunjukkan 1 hidden layer dengan sizes 9 dan 2 hidden layer dengan size 4 menghasilkan nilai akurasi tertinggi yaitu 70.98%.

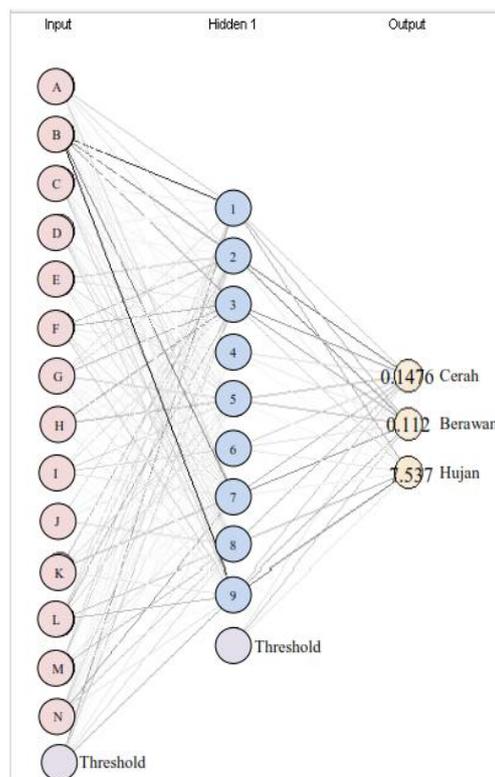
Berdasarkan eksperimen yang telah dilakukan, didapatkan hasil arsitektur

yang dianggap terbaik yaitu dengan nilai akurasi yang paling tinggi.

Tabel 9. Parameter Neural Network hasil eksperimen

Training Cycle	Learning Rate	Momentum	Σ	H 2	Accuracy
500	0.1	0.1	1	9	72.97%

Arsitektur jaringan yang terbaik dari hasil eksperimen dapat dilihat pada Gambar 3.



Gambar 3. Arsitektur Jaringan yang didapat dari Hasil Eksperimen

Tabel 10. Nilai bobot akhir untuk input layer

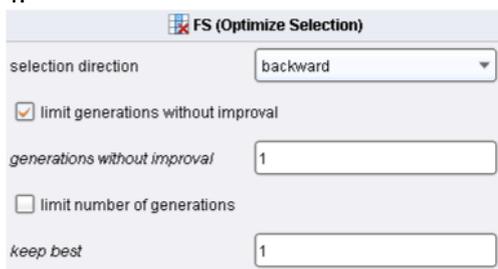
Node	Hidden Layer (Sigmoid)														
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	Threshold
1	2.756	10.284	0.786	0.791	0.317	0.035	-	-	-	0.154	0.217	1.851	-	-	-5.207
2	2.009	7.005	0.675	0.665	2.803	3.670	-	2.108	0.746	0.146	-	-	2.345	0.891	0.806
3	-	5.575	0.816	0.843	-	3.957	-	2.308	0.505	-	2.156	3.893	2.357	-	0.806
4	-	-0.311	0.251	0.172	-	1.051	0.252	4.198	5.603	2.138	0.206	0.370	4.475	0.668	-2.782
5	0.704	-0.458	1.638	1.628	2.055	-	-	4.201	-	-	-	-	0.214	0.253	-0.089
6	-	1.309	0.346	0.330	-	0.334	0.458	0.026	2.019	-	1.517	-	-	-	-0.058
7	2.318	-7.553	-	-	-	-	-	-	-	-	-	3.835	-	-	1.268
8	0.904	0.563	1.815	1.860	-	1.391	-	-	-	-	0.030	0.720	1.738	-	-2.259
9	-	-	-	-	-	2.880	-	1.216	0.614	0.387	1.189	3.163	-	1.488	4.390

Tabel 3.11. Nilai bobot akhir untuk Output layer

Class	Output Sigmoid									
	1	2	3	4	5	6	7	8	9	Threshold
Cerah	2.768	-7.620	5.451	-1.632	-3.707	-2.431	3.981	-4.038	3.379	-1.450
Berawan	-4.806	7.700	-5.076	-1.202	3.671	1.448	-5.536	0.135	4.650	-3.391
Hujan	4.596	1.215	2.468	-0.562	-0.619	-1.097	0.764	3.888	-7.415	-1.456

3.1.3. Penerapan Backward Elimination Pada Neural Network

Dalam eksperimen Neural Network berbasis backward elimination spesifikasi pada backward elimination pada Gambar 4.

**Gambar 4.** Spesifikasi Backward Elimination

3.1.3.1. Parameter Neural Network berbasis Backward Elimination

Eksperimen pertama dilakukan untuk mendapatkan nilai training cycles, yaitu dilakukan uji coba dengan memasukan nilai dengan range 100 sampai dengan 700, nilai 0.1 untuk learning rate dan nilai 0.1 untuk momentum. Berikut hasil eksperimen terhadap training cycles:

Tabel 12. eksperimen untuk menentukan nilai training cycles

Training Cycle	Learning Rate	Momentum	Accuracy
100	0.1	0.1	74.72%
200	0.1	0.1	73.21%
300	0.1	0.1	72.94%
400	0.1	0.1	75.19%
500	0.1	0.1	74.70%
600	0.1	0.1	74.70%
700	0.1	0.1	70.74%

Hasil eksperimen menunjukkan bahwa nilai training cycles dengan accuracy tertinggi terdapat pada training cycles dengan nilai 400 dengan accuracy 75.19 %.

Tabel 13. eksperimen untuk menentukan nilai learning rate

Training Cycle	Learning Rate	Momentum	Accuracy
400	0.1	0.1	74.72%

400	0.2	0.1	74.16%
400	0.3	0.1	74.68%
400	0.4	0.1	75.91%
400	0.5	0.1	72.21%

Dari hasil eksperimen diatas didapat bahwa nilai learning rate dengan Accuracy tertinggi yaitu pada learning rate dengan nilai 0.4 dengan Accuracy 75.91%

Tabel 14. eksperimen untuk menentukan nilai learning rate

Training Cycle	Learning Rate	Momentum	Accuracy
400	0.4	0.1	75.91%
400	0.4	0.2	75.45%
400	0.4	0.3	71.97%
400	0.4	0.4	72.02%
400	0.4	0.5	72.93%

Dari hasil eksperimen diatas didapat bahwa nilai momentum dengan Accuracy tertinggi yaitu pada momentum dengan nilai 0.1 dengan Accuracy 75.91%

3.1.3.2. Menentukan Jumlah Hidden Layer dan Hidden Layer Sizes

Pada jumlah hidden layer sebanyak 1, dilakukan percobaan pada hidden layer sizes dengan range 1 sampai dengan 3 sizes. Berikut hasil eksperimen dengan 1 hidden layer :

Tabel 15. Eksperimen untuk menentukan Hidden Layer Sizes pada Hidden Layer 1

Training Cycle	Learning Rate	Momentum	Sizes	Accuracy
400	0.4	0.1	1	73.23%
400	0.4	0.1	2	74.45%
400	0.4	0.1	3	73.23%

Hasil eksperimen menunjukkan 1 hidden layer dengan sizes 3 yang menghasilkan akurasi tertinggi yaitu, dengan nilai 74.45%.

Tabel 16. Eksperimen untuk menentukan Hidden Layer Sizes pada Hidden Layer 2

Training Cycle	Learning Rate	Momentum	H1	H 2	Accuracy
400	0.4	0.1	2	1	64.25%
400	0.4	0.1	2	2	76.94%
400	0.4	0.1	2	3	74.98%

Hasil eksperimen menunjukkan 2 hidden layer dengan sizes 2 yang menghasilkan akurasi tertinggi yaitu, dengan nilai 76.94%.

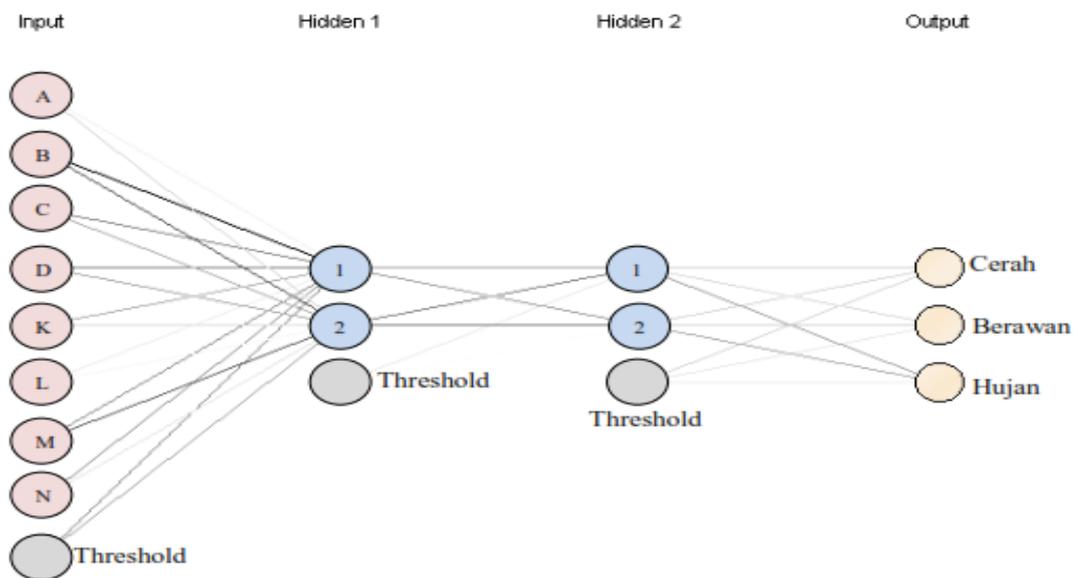
Dari hasil eksperimen diatas didapat arsitektur jaringan yang dianggap terbaik yaitu dengan nilai accuracy paling tinggi, tampak pada Tabel 17.

Terakhir adalah implementasi dataset hasil Backward Selection kedalam metode Neural Network, hal ini guna mendapatkan arsitektur jaringan yang

terbaik. Setelah data di running maka, telah didapatkan arsitektur jaringan yang terbaik dari hasil eksperimen tampak pada Gambar 5.

Tabel 17. Parameter Neural Network berbasis Backward Elimination hasil eksperimen

Training Cycle	Learning Rate	Momentum	H1	H2	Accuracy
400	0.4	0.1	2	2	76.94%



Gambar 5. Arsitektur Jaringan yang didapat dari Hasil Eksperimen Dengan Backward Elimination

Berdasarkan gambar 5. model yang didapat adalah model dengan pola 9- 3-3-3 dengan accuracy 76.94%. Hasil eksperimen yang telah dilakukan dengan

metode Neural Network dan metode Neural Network berbasis Backward Selection maka didapatkan hasil seperti pada Tabel 18.

Tabel 18. Komparasi Accuracy

Method	Training Cycle	Learning Rate	Momentum	H1	H2	Accuracy
NN+BE	400	0.4	0.1	2	2	76.94%
NN	500	0.4	0.1	2	2	72.97%

Berdasarkan tabel 3.18, Hasil penelitian menunjukkan metode jaringan syaraf tiruan berbasis backward elimination menghasilkan peningkatan akurasi 4% dibandingkan hanya dengan menggunakan metode jaringan syaraf tiruan saja. Akurasi yang dihasilkan dengan metode Neural Network berbasis Backward Elimination dapat meningkatkan hasil akurasi lebih tinggi yaitu 76.94% jika dibandingkan dengan Akurasi yang dihasilkan dengan metode Neural Network

yaitu 72.97%. Hal ini menunjukkan bahwa prediksi cuaca dengan menggunakan Neural Network berbasis Backward Elimination lebih akurat. Melalui penlitian ini dapat diketahui sejauh mana metode Neural Network menggunakan seleksi fitur Backward Elimination mempunyai tingkat akurasi lebih baik sehingga dapat membantu banyak pihak yang membutuhkan informasi kondisi atmosfer (cuaca) yang lebih cepat, lengkap, dan akurat.

4. KESIMPULAN

Dari penelitian yang dilakukan model yang terbentuk dengan algoritma Neural Network sendiri memiliki tingkat akurasi yaitu sebesar 72.97% dengan proses seleksi fitur oleh Backward Elimination, model yang terbentuk dapat meningkatkan akurasi menjadi 76.94%. Hasil penelitian menunjukkan metode jaringan syaraf tiruan berbasis backward elimination menghasilkan peningkatan akurasi 4% dibandingkan hanya dengan menggunakan metode Neural Network saja.

Dengan model prediksi yang baik, dapat memberikan informasi bagi kebijakan pemerintah dalam perencanaan pengelolaan sumber daya air kedepannya dan juga dapat memberikan informasi cuaca sedini mungkin bagi masyarakat pengguna jasa agar dapat melakukan aktivitas kegiatannya untuk beberapa waktu kedepan.

Proses seleksi fitur dengan menggunakan algoritma Backward Elimination terbukti dapat meningkatkan akurasi dari algoritma Neural Network. Namun ada beberapa faktor yang dapat dicoba untuk penelitian selanjutnya.

Agar algoritma Neural Network dapat menghasilkan model yang lebih baik lagi, sebaiknya menggunakan attribute lebih sedikit tetapi juga menggunakan jumlah data training lebih banyak lagi, pengurangan waktu komputasi untuk dataset yang besar.

Hasil penelitian ini diharapkan dapat diterapkan pada salah satu metode manajemen penanggulangan banjir akibat perubahan cuaca, dengan melihat intensitas terjadinya hujan.

DAFTAR PUSTAKA

- Aldrian, E., & Djamil, Y. S. (2008). Application of Multivariate ANFIS For Daily Rainfall Prediction: Influences of Training Data Size. *MAKARA*, 12(April 2008), 7–14.
- Beniwal, S., & Arora, J. (2012). Classification and Feature Selection Techniques in Data Mining. *International Journal of Engineering Research & Technology (IJERT)*, 1(6), 1–6.
- Chris Fraley, Adrian Raftery, Tilmann Gneiting, McLean Slaughter & Veronica Berrocal. Probabilistic Weather Forecasting in R. *The R Journal*. 56-62.
- Felkin, M. (2007). between N -ary and Binary Problems. *Between N-ary and Binary Problems*, 1–25.
- Han & Kamber. (2006). *Data Mining: Concept and Technique* (2nd ed). United State America. Heaton.
- (2008). *Introduction to Neural Network With Java* (2nd ed). USA. Heaton Research, Inc.
- Hong, X., Harris, C., Brown, M., & Chen, S. (2002). Backward Elimination Methods for Associative Memory Network Pruning. *Computers and Technology*, (Reed 1993).
- Hung, N. Q., Babel, M. S., Weesakul, S., & Tripathi, N. K. (2009). An artificial neural network model for rainfall forecasting in Bangkok , Thailand. *Hydrology and Earth System Sciences*, 1413–1425.
- Jong Jek Siang. (2009). *Jaringan Syaraf Tiruan & Pemrogramannya Menggunakan MATLAB*. Yogyakarta : Andi Yogyakarta.
- Kohavi, R., & John, H. (1997). Artificial Intelligence Wrappers for feature subset selection. *elsevier*, 97(97), 273–324.
- Kusumadewi, Sri & Hartati, Sri. (2010). *Neuro-Fuzzy Integrasi Sistem Fuzzy & Jaringan Syaraf*. (2nd ed). Yogyakarta: Graha Ilmu.
- Lin, S.-W., Chen, S.-C., Wu, W.-J., & Chen, C.-H. (2009). Parameter determination and feature selection for back-propagation network by particle swarm optimization. *Knowledge and Information Systems*, 21(2), 249–266.
- Liu, H., Member, S., Yu, L., & Member, S. (2005). Algorithms for Classification and Clustering. *IEEE Transaction on Knowledge and Data Engineering*, 17(April 2005), 491–502. Program Pascasarjana Magister Ilmu Komputer STMIK Nusa Mandiri

-
- Pallavi, V. P., Vaithyanathan, V., & Ph, D. (2013). Combined Artificial Neural Network and Genetic Algorithm for Cloud Classification. *International Journal of Engineering Research & Technology (IJET)*, (May), 787–794.
- Radhika, Y., & Shashi, M. (2009). Atmospheric Temperature Prediction using Support Vector Machines. *International Journal Of Computer Theory and Engineering*, 1(April), 55–58. doi:10.7763/IJCTE.2009.V1.9
- Selection, O. F., & Elimination, B. (2004). Orthogonal Forward Selection and Backward Elimination for Feature Subset Selection. *IEEE Transactions on System Man and Cybernetics*, 1(February), 629–634.
- Solaimani, K. (2009). Rainfall-runoff Prediction Based on Artificial Neural Network (A Case Study : Jarahi Watershed). *IDOSI Publication*, 5(6), 856– 865.
- Vercellis, C. (2009). *Business Intelligence: Data Mining and Optimization for Decision Making*. Wiley.
- Winarso. (2002). *Pemikiran dan Praktek Perencanaan dalam Era Tranformasi di Indonesia*. Bandung: Departemen Teknik Planologi ITB.
- Zhu, Z., Ong, Y., & Dash, M. (n.d.). Wrapper-Filter Feature Selection Algorithm Using A Memetic Framework. *International Journal of Engineering Research & Technology (IJET)*, 1–19.