

Hibridisasi Algoritma *Classification and Regression Tree (CART)* dan *Artificial Neural Network (ANN)* untuk Prediksi Harga Bawang Merah di Kabupaten Bandung

Himatul Zulfa¹, Fhira Nhita², Deni Saepudin³

^{1,2,3}Fakultas Informatika, Universitas Telkom, Bandung

¹himatulzulfa@students.telkomuniversity.ac.id, ²fhiranhita@telkomuniversity.ac.id,

³denisaepudin@telkomuniversity.ac.id

Abstrak

Bawang merah adalah salah satu jenis tumbuhan hortikultura yang mempunyai banyak manfaat namun kualitas dan kuantitas nya sangat dipengaruhi oleh cuaca. Perubahan cuaca akan menyebabkan produksinya tidak menentu sehingga mengalami fluktuasi harga. Prediksi harga terhadap cuaca ini sangat diperlukan agar mempermudah petani untuk melakukan penanaman pada waktu yang tepat. Oleh sebab itu, penelitian ini memprediksi harga bawang merah di Kabupaten Bandung berdasarkan pengaruh cuaca dengan menggunakan hibridisasi dari *Classification and Regression Tree (CART)* dan *Artificial Neural Network (ANN)*. Berdasarkan beberapa skenario pengujian dalam penelitian ini, metode hibridisasi CART-ANN terbukti memberikan hasil yang lebih baik daripada CART atau ANN secara konvensional dengan nilai *precision*, *recall*, dan akurasi berturut-turut sebesar 90.91%, 100.00%, 93.33%.

Kata kunci : Bawang Merah, *Classification and Regression Tree*, *Artificial Neural Network*

Abstract

Shallot is one type of horticultural plant that has many benefits, but the weather greatly influences its quality and quantity. Weather changes will cause production to be erratic so that prices fluctuate. Price forecast for weather is essential to make it easier for farmers to plant at the right time. Therefore, this study predicts the price of shallots in Bandung Regency based on the influence of weather using hybridization from Classification and regression tree (CART) and Artificial Neural Network (ANN). Based on several test scenarios in this study, the CART-ANN hybridization method proved to provide better results than conventional CART or ANN with the values of precision, recall, and accuracy contributing respectively 90.91%, 100.00%, 93.33%.

Keywords: Shallot, *Classification and Regression Tree*, *Artificial Neural Network*

1. Pendahuluan

Pendahuluan ini terdiri dari empat sub-bagian yaitu: Latar Belakang, Topik dan Batasannya, Tujuan dan Organisasi Tulisan. Berikut penjelasan dari masing-masing sub-bagian tersebut.

Latar Belakang

Bawang merah (*Allium ascalonicum* L.) merupakan salah satu tumbuhan komoditas hortikultura yang penting dan banyak dibutuhkan oleh masyarakat [1]. Bawang merah memiliki nilai ekonomi yang tinggi di Indonesia, dan produksinya pun di Indonesia menempati area budidaya terbesar ketiga, setelah cabai dan kubis [2]. Namun meskipun demikian, pasokan bawang merah sering menjadi masalah, karena kebutuhannya terus meningkat dari tahun ketahun, hingga mencapai 4,6 kg per kapita per tahun, sedangkan produksi bawang merah tidak menentu mengakibatkan harganya terus mengalami fluktuasi [3].

Pasokan bawang merah di pertanian domestik tidak menentu karena dipengaruhi oleh beberapa faktor, diantaranya yaitu: faktor pemilihan varietas benih, evaluasi kesesuaian lahan dan faktor cuaca. Namun, dari ketiga faktor tersebut, faktor cuaca merupakan salah satu faktor utama yang mempengaruhi produksi bawang merah [4]. Pada pulau Jawa pada tahun 2013 terjadi perubahan cuaca yang ekstrim sehingga menyebabkan produksi bawang merah menurun dan harga bawang merah melonjak hingga Rp60.000,00/ kg [5]. Bahkan pada kasus lain, Jika cuaca buruk hasil produksi bawang merah bisa menjadi sangat rendah, yaitu 50% dari hasil normal [6]. Oleh karena itu, prediksi harga bawang merah yang dipengaruhi oleh cuaca ini sangat perlu dilakukan agar mempermudah petani untuk melakukan penanaman pada waktu yang tepat dengan lebih mempertimbangkan faktor cuaca.

Penelitian tentang prediksi harga komoditas pertanian sebelumnya sudah pernah dilakukan, seperti pada penelitian *Planting Date Recommendation for Chili and Tomato Based on Economic Value Prediction of Agricultural Commodities* menggunakan algoritma ANN-NGA menghasilkan akurasi masing-masing 81% dan 100% [7], pada penelitian *Price Prediction of Chili Commodities in Bandung Regency* dengan menggunakan algoritma *Bayesian Network* menghasilkan akurasi rata-rata 83,5% dalam mengklasifikasikan harga [8], dan pada penelitian *Price Prediction for Chili in Bandung Regency* menggunakan algoritma SVM yang dioptimalkan oleh ANFIS menghasilkan keakuratan 94,07% [9]. Namun harga komoditas barang merah belum pernah diprediksi sebelumnya dan metode hibridisasi dari CART-ANN sendiri juga belum pernah ditemukan untuk memprediksi harga komoditas pertanian. Metode ini pernah digunakan untuk memprediksi potensi penempatan seorang siswa. Hasil yang diperoleh sangat meningkat dari pada hanya menggunakan CART atau ANN saja [10]. Hibridisasi CART-ANN berpotensi akan memberikan hasil akurasi yang lebih tinggi. Oleh karena itu, penelitian ini mengajukan model hibridisasi algoritma CART-ANN untuk memprediksi harga bawang merah yang dipengaruhi oleh cuaca agar mendapatkan hasil akurasi yang lebih baik.

Topik dan Batasannya

Kebutuhan bawang merah terus meningkat setiap tahun, sedangkan produksi bawang merah tidak menentu mengakibatkan harganya terus mengalami fluktuasi. Pada penelitian ini penulis mengangkat judul tentang pengaruh cuaca terhadap harga bawang merah khususnya di Kabupaten Bandung dengan mencari atribut penting dari semua atribut yang ada dengan menggunakan algoritma CART kemudian digabungkan/ dihibridisasi dengan algoritma ANN untuk memperoleh kinerja yang lebih baik. Batasan data yang digunakan yaitu dari data bulanan cuaca, dan harga bawang merah dari tahun 2014 hingga 2017.

Tujuan

Penelitian ini bertujuan untuk menganalisa pengaruh perubahan cuaca terhadap harga bawang merah khususnya di Kabupaten Bandung, memilih atribut penting/ yang paling berpengaruh dari semua atribut yang ada dengan menggunakan model CART dan membuktikan bahwa model hibridisasi CART-ANN memberikan kinerja sistem yang lebih baik daripada model CART atau ANN secara konvensional.

Organisasi Tulisan

Selanjutnya pada jurnal TA ini akan menjelaskan literatur yang mendukung dengan penelitian ini pada bagian 2. Penjelasan terkait sistem yang akan dibangun terdapat pada bagian 3. Bagian 4 menjelaskan hasil dan evaluasi yang telah dilakukan. Terakhir yaitu kesimpulan dan saran untuk penelitian kedepannya dijelaskan pada bagian 5.

2. Studi Terkait

2.1. Algoritma *Classification and Regression Tree* (CART)

CART diperkenalkan oleh Breiman et al yang didasarkan pada dua variabel, yaitu variabel dependen kualitatif (klasifikasi) dan variabel kuantitatif (pohon regresi) [11]. CART adalah salah satu algoritma *Decision Tree* (DT) yang banyak digunakan karena dapat memberikan tafsiran sederhana dengan menggunakan hierarki struktur pohon untuk tujuan klasifikasi dan regresi [12]. Tujuan dari CART adalah untuk mendapatkan klasifikasi optimal dan atribut penting agar prediksi harga terhadap cuaca yang dihasilkan lebih akurat. Berikut langkah-langkah algoritma CART pada penelitian ini:

1. *Split process* adalah proses pemecahan node yang diawali dengan menentukan variabel pemecah terbaik dengan menggunakan indeks gini terendah. Atribut yang memiliki indeks gini terendah maka akan menjadi *node root*. Berikut cara menghitung gini indeks [13]:

$$Gini(D) = 1 - \sum_{i=1}^N p_i^2 \quad (1)$$

Dengan :

p_i : *probability* kelas i , i adalah kelas positif dan negatif.

N : Nomor label kelas

Jika D dibagi menjadi D_1 dan D_2 berdasarkan atribut A , maka

$$Gini_{Split(A)}(D) = \frac{D_1}{D} Gini(D_1) + \frac{D_2}{D} Gini(D_2) \quad (2)$$

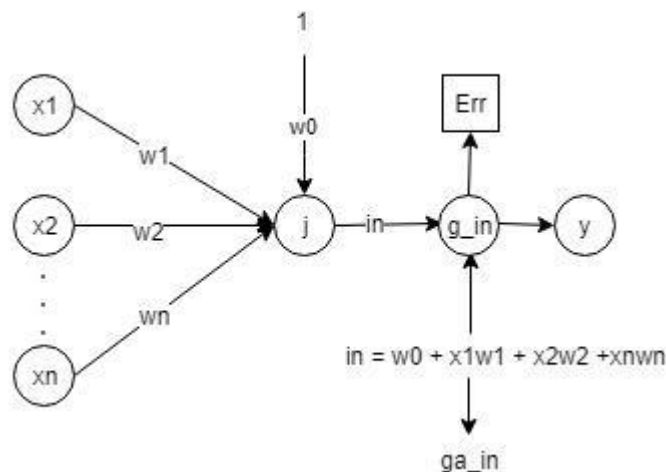
Dengan : Gini D_1 dan D_2 dihitung dari rumus ke (1)

2. *Partition process* adalah proses mengulang *split* proses sampai mencapai kondisi berhenti yang telah ditentukan. Kondisi berhenti pada penelitian ini yaitu dengan menentukan kedalaman pohon maximum. Jika penggunaan nilai kedalaman pohon maximum nya terlalu tinggi maka akan menghasilkan model yang kompleksitas nya tinggi, sedangkan jika terlalu rendah maka akan menyebabkan penghapusan beberapa parameter input [11]. Oleh karena itu perlu dilakukannya *trial* dan *error* untuk mendapatkan kedalaman pohon maximum yang sesuai.
3. *Pruning process* adalah sebuah proses pemangkasan yang dapat mengurangi data *outlier* dan *noisy* sehingga dapat meningkatkan akurasi dalam klasifikasi data [14].
4. *Rule extraction process* adalah pengestraksi aturan if-then

Kemudian setelah melakukan semua proses diatas, maka CART akan memberikan rekomendasi atribut terpenting dan hasil prediksi CART. Atribut penting dan hasil prediksi CART ini nantinya akan dijadikan input *node* pada proses hibridisasi CART-ANN.

2.2. Algoritma Artificial Neural Network (ANN)

Artificial Neural Network (ANN) adalah suatu jaringan yang memodelkan sistem saraf otak manusia (yang disebut neuron) dalam melakukan pengenalan pola, khususnya klasifikasi [15]. Salah satu algoritma pembelajaran pada ANN yang sederhana adalah perceptron dengan lapisan tunggal. Perceptron diperkenalkan oleh Frank Rosenblatt dengan aturan pembelajaran menggunakan sebuah pengaturan bobot *iterative* yang lebih baik dibandingkan dengan aturan pembelajaran Hebb [16]. Arsitektur perceptron belajar mengenali pola dengan metode belajar terbimbing. Teknik ini hanya mempunyai sebuah lapisan input dan sebuah unit output seperti pada gambar 1.



Gambar 1. Model Perceptron

Pada model tersebut sinyal masukan (*input*) dinotasikan sebagai $x_1, x_2, \dots, x_n$, dan masukan bias bernilai tetap yaitu 1 (*fixed input*). Sedangkan bobot-bobot dinotasikan sebagai $w_0, w_1, w_2, \dots, w_n$. Kemudian unit pengolah (j) melakukan penjumlahan berbobot untuk seluruh masukannya, dengan rumus sebagai berikut:

$$in_i = w_0 + \sum_{i=1}^n w_i x_i \quad (3)$$

Hasil in_i ini kemudian akan digunakan untuk perhitungan g_in dan ga_in dimana g merupakan fungsi aktivasi. Fungsi aktivasi merupakan bagian penting dalam tahapan perhitungan keluaran dari suatu algoritma [17]. Fungsi aktivasi yang digunakan pada penelitian ini yaitu fungsi aktivasi sigmoid. Fungsi ini bisa digunakan untuk jaringan saraf yang nilai outputnya 0 atau 1.

$$g_in_i = \frac{1}{1 + e^{-in_i}} \quad (4)$$

Setelah itu dilakukan perhitungan *error* dengan selisih nilai g_in_i dengan kelas target. Pada perceptron proses pembelajaran untuk mendapatkan bobot-bobot akhir dilakukan secara berulang sampai sudah tidak terjadi kesalahan (*error*). Apabila proses pembelajaran sudah dilakukan sampai berulang-ulang, namun masih saja terjadi kesalahan, maka proses pembelajaran harus dihentikan. Sehingga diperlukan adanya pembatasan lama perulangan, yang sering dikenal dengan nama maksimum epoch. Selain itu, dalam melakukan pembaruan

bobot, proses pembelajaran dengan perceptron juga dipengaruhi oleh *learning rate* yang disimbolkan dengan α bernilai 0 sampai 1, semakin mendekati 1 maka proses pembelajaran semakin cepat [18]. Formula yang digunakan untuk pembaruan bobot baru adalah:

$$ga_in_i = (1 - g_in_i) \quad (5)$$

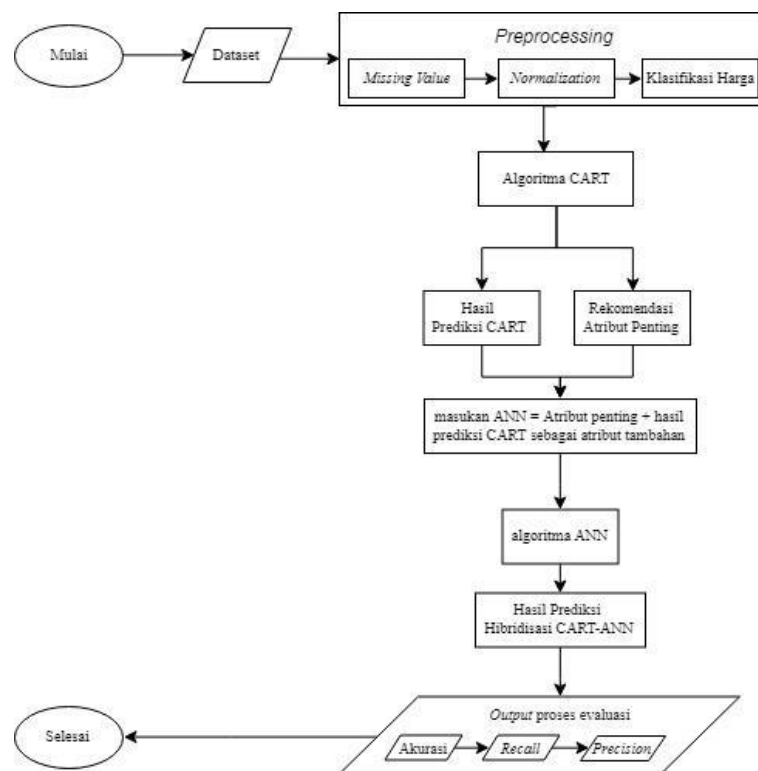
$$w_{0baru} = (\alpha * error_i * ga_in_i * b) \quad (6)$$

$$w_{ibaru} = (\alpha * error_i * ga_in_i * x_i) \quad (7)$$

Bobot-bobot ini akan terus dilakukan pembaruan sampai tidak ada lagi bobot yang berubah atau sampai mencapai maksimum epoch.

3. Sistem yang Dibangun

Pada bagian ini, akan dijelaskan desain sistem dari hibridisasi algoritma CART-ANN untuk memprediksi harga bawang merah yang dipengaruhi oleh cuaca.



Gambar 2. Gambaran Umum Sistem

3.1. Dataset

Dataset yang digunakan pada penelitian ini merupakan data harga bawang merah dari Dinas Perindustrian dan Perdagangan Kabupaten Bandung dan data cuaca dari Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) Kabupaten Bandung dari tahun 2014-2017. Data yang digunakan sebagai data inputan terdapat tujuh atribut, diantaranya yaitu: lama penyinaran matahari (PM), kecepatan angin (KA), temperatur (TM), kelembapan (KL), penguapan (PU), curah hujan (CH) dan harga pasar actual (HPA).

3.2. Preprocessing

Preprocessing data merupakan langkah yang paling penting dalam proses Data Mining [19] untuk mendapatkan akurasi tertinggi [14]. Berikut beberapa teknik *Preprocessing* data yang digunakan pada penelitian ini agar menghasilkan data yang lebih baik:

1. Missing Value

Missing value adalah informasi yang hilang atau tidak tersedia untuk sebuah objek (kasus). Penanganan *missing value* bisa dengan dua cara, yaitu: mean (rata-rata) dan modus (nilai yang sering muncul) dari atribut tersebut untuk mengisi nilai yang hilang [14].

2. Normalization

Normalisasi adalah proses yang menghasilkan range nilai dari *interval* 0 sampai 1 yang bertujuan meminimalkan skor kesalahan [20].

$$X'_n = \frac{X_n - Min_x}{Max_x - Min_x} \quad (8)$$

Dimana :

X'_n = setelah normalisasi data (hasil)

X_n = sebelum normalisasi data

Min_x = minimum data pada variabel x

Max_x = maximum data pada variabel x

3. Klasifikasi Harga

Pengklasifikasian harga dibagi menjadi dua yaitu kelas ekonomi dan kelas tidak ekonomi dinilai berdasarkan dari perbandingan harga jual petani dengan nilai biaya produksi. Harga jual petani yaitu 75% dari harga jual pasar. Berdasarkan informasi dari petani sayuran di Kabupaten Bandung diperoleh harga produksi bawang merah pada bulan Juli 2017 sebesar Rp 20,000. Setelah mengetahui harga produksi pada bulan Juli 2017 maka digunakan rumus *Future Value (Fv)* untuk mengetahui harga biaya produksi pada bulan sebelum dan sesudahnya [7,8].

$$Fv = Pc \times e^{-it} \quad (9)$$

Dimana :

Fv = future value biaya produksi per kilogram

Pc = biaya produksi perkilogram bulan ini

i = rata – rata inflasi perbulan

t = waktu

Dari perhitungan diatas maka akan diperoleh biaya produksi selama 4 tahun. Langkah selanjutnya adalah membandingkan harga jual petani dengan biaya produksi untuk mendapatkan kelas harga, sebagaimana tabel berikut:

Tabel 1. Klasifikasi Harga

Perbandingan	Kelas
Harga jual petani $\leq$ Biaya produksi	Tidak Ekonomis (0)
Harga jual petani $>$ Biaya produksi	Ekonomis (1)

3.3. Proses Hibridisasi CART-ANN

CART merupakan model klasifikasi yang unggul dari pada ANN dalam membaca informasi, sedangkan ANN adalah model yang lebih unggul dalam inferensi komprehensif atas *input* [10]. Oleh karena itu, kami mengusulkan penggabungan/ hibridisasi CART dengan klasifikasi lain yaitu ANN agar hasil yang didapatkan lebih baik. Pada penelitian yang merujuk pada referensi [10] penggabungan CART-ANN memberikan hasil yang lebih baik daripada CART atau ANN secara konvensional.

Tabel 2. Metodologi Hibridisasi CART-ANN

Algoritma Desain Sistem
Masukan : Data cuaca dan harga bawang merah yang sudah dinormalisasi (x), dan kelas (y)
Keluaran : Evaluasi performansi dan kelas hasil prediksi
Algoritma:
1 : Melatih dan membangun CART untuk mendapatkan atribut-atribut penting
2 : Hasil prediksi CART digunakan sebagai atribut tambahan pada lapisan masukan ANN
3 : Masukkan atribut-atribut penting dari CART bersama variabel atribut tambahan ke lapisan input ANN.
4 : Jalankan ANN hingga akurasi yang dihasilkan baik
Analisis Performansi:
5: Mengukur ketepatan klasifikasi dengan evaluasi performansi.

Hasil prediksi dari CART dimasukkan sebagai fitur input adalah untuk meningkatkan dimensi ruang fitur dan juga akan meningkatkan pemisahan kelas, karena jika ruang fitur tidak padat maka ruang fitur akan terpisah secara linear. Selain itu, kinerja CART-ANN akan meningkat apabila semakin banyak informasi dimasukkan [10].

3.4. Evaluasi Performansi

Evaluasi performansi digunakan untuk mengukur kinerja algoritma klasifikasi yaitu seberapa bagus atau seberapa akurat performansi sistem yang telah dibuat dengan membandingkan hasil klasifikasi yang dilakukan oleh sistem dengan hasil klasifikasi yang sebenarnya [21]. Evaluasi performansi yang digunakan adalah *Confusion matrix*. Berikut rumus yang digunakan untuk mengukur kinerja algoritma klasifikasi pada *Confusion matrix* [21]:

$$1. \text{ Accuracy} = \frac{TP + TN}{TP + FN + FP + TN} \quad (10)$$

Accuracy adalah pengukuran yang menggambarkan seberapa akurat sistem dalam mengklasifikasikan data secara benar.

$$2. \text{ Recall} = \frac{TP}{TP + FN} \quad (11)$$

Recall adalah untuk mengetahui tingkat keberhasilan sistem dalam mengenali suatu data dari seluruh data yang seharusnya dikenali.

$$3. \text{ Precision} = \frac{TP}{TP + FP} \quad (12)$$

Precision adalah pengukuran yang menggambarkan tingkat ketepatan sistem.

Tabel 3. *Confusion Matrix*

Kelas Prediksi			
Sebenarnya		Tidak ekonomis	ekonomis
	Tidak ekonomis	TN	FP
	Ekonomis	FN	TP

Berikut penjelasan dari tabel diatas:

1. TN adalah kelas tidak ekonomis yang terprediksi tidak ekonomis dan menjadi jumlah nilai tidak ekonomis yang benar.
2. FP adalah kelas tidak ekonomis yang terprediksi ekonomis dan menjadi jumlah nilai yang terprediksi ekonomis.
3. FN adalah kelas ekonomis yang terprediksi tidak ekonomis dan menjadi jumlah nilai yang terprediksi tidak ekonomis.
4. TP adalah kelas ekonomis yang terprediksi ekonomis dan menjadi jumlah nilai kelas ekonomis yang benar.

4. Evaluasi

Bagian ini terdiri dari dua sub-bagian yaitu hasil dan analisis dari pengujian yang telah dilakukan.

4.1 Hasil Pengujian

Hal pertama yang dilakukan adalah membuat skenario pengujian. Skenario pengujian dilakukan dengan mempartisi data menjadi lima kondisi partisi yang berbeda. Berikut tabel pembagian partisi data yang dibagi menjadi data *training* dan data *testing*.

Tabel 4. Skenario Partisi Data

Data Partisi	Data Training	Jumlah Data Training	Data Testing	Jumlah Data Testing
1	Jan '14 - Des '15	24	Jan '16 - Nov '17	23
2	Jan '14 - Apr '16	28	May '16 - Nov '17	19
3	Jan '14 - Aug '16	32	Sept '16 - Nov '17	15
4	Jan '14 - Dec '16	36	Jan '17 - Nov '17	11
5	Jan '14 - Apr '17	40	May '17 - Nov '17	7

Setelah dilakukan partisi data, kemudian dilakukan perincian jumlah kelas ekonomis (1) dan tidak ekonomis (0) pada setiap partisi data pada tabel 4. sebagaimana tabel dibawah ini.

Tabel 5. Jumlah Kelas pada Partisi

Kelas	Partisi 1		Partisi 2		Partisi 3		Partisi 4		Partisi 5	
	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test
ekonomis (1)	5	17	8	14	12	10	16	6	19	3
tidak ekonomis (0)	19	6	20	5	20	5	20	5	21	4

Kemudian setiap partisi data tersebut akan diproses menggunakan algoritma CART pada python 3.6 untuk mendapatkan hasil atribut terpenting. Berikut tabel hasil dari proses CART yang menghasilkan atribut terpenting.

Tabel 6. Atribut Terpenting CART

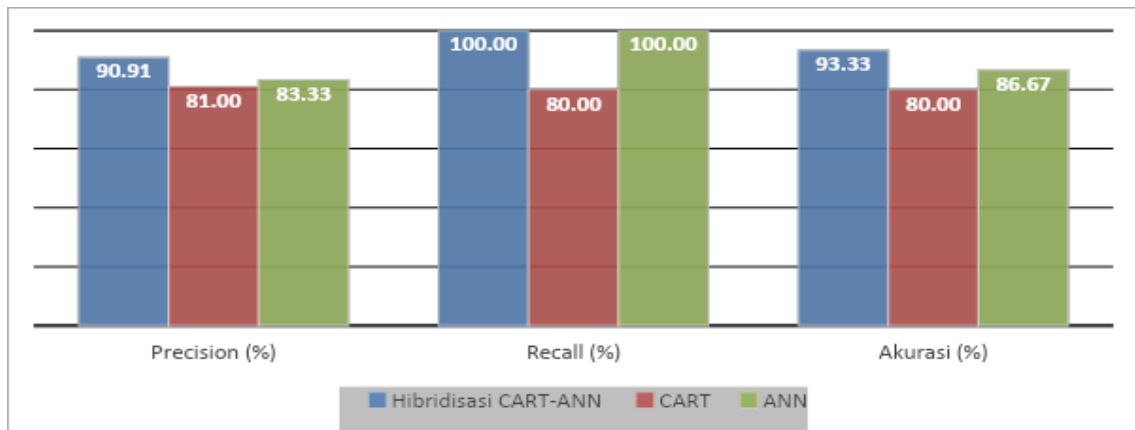
Kedalaman pohon	Partisi 1	Partisi 2	Partisi 3	Partisi 4	Partisi 5
1	HPA	HPA	HPA	HPA	HPA
2	CH, HPA	CH, HPA	PU, CH, HPA	CH, HPA	CH, HPA
3	TM, CH, HPA	PU, CH, HPA	TM, PU, CH, HPA	TM, PU, CH, HPA	TM, PU, CH, HPA
4	TM, CH, HPA	PU, CH, HPA	TM, PU, CH, HPA	TM, PU, CH, HPA	TM, PU, CH, HPA

Sebagaimana yang dapat dilihat pada tabel 6. bahwa kedalaman pohon 1 untuk semua partisi hanya menghasilkan satu atribut penting yaitu HPA dan untuk kedalaman pohon 3-4 menghasilkan atribut penting yang sama. Pada proses hibridisasi CART-ANN atribut yang digunakan minimal harus dua atribut yaitu atribut harga dan atribut cuaca, khususnya curah hujan. Sehingga yang di proses lebih lanjut ke tahap hibridisasi CART-ANN yaitu untuk kedalaman pohon 2 dan 3 disetiap partisi data. Kemudian, atribut penting dan hasil prediksi dari CART dijadikan sebagai inputan pada model ANN. Model ANN pada penelitian ini diproses menggunakan MATLAB R2017a. Berikut tabel hasil dari proses hibridisasi CART-ANN.

Tabel 7. Hibridisasi CART-ANN

Skenario	Precision (%)	Recall (%)	Accuracy (%)	Confusion Matrix
1_2	73.91	100.00	73.91	0 6 0 17
1_3	76.19	94.12	73.91	1 5 1 16
2_2	80.00	57.14	57.89	3 2 6 8
2_3	85.71	85.71	78.95	3 2 2 12
3_2	90.91	100.00	93.33	4 1 0 10
3_3	80.00	80.00	73.33	3 2 2 8
4_2	71.43	83.33	72.73	3 2 1 5
4_3	60.00	100.00	63.64	1 4 0 6
5_2	50.00	66.67	57.14	2 2 1 2
5_3	42.86	100.00	42.86	0 4 0 3

Hasil pengujian dari tiap-tiap skenario diatas menunjukkan bahwa hibridisasi CART-ANN khususnya pada skenario ke 3_2 (skenario ke tiga dengan menggunakan nilai kedalaman dua) memberikan kinerja sistem yang lebih baik daripada skenario lainnya karena dari 15 data pengujian yang diberikan, hanya satu kelas yang tidak dapat dikenali. skenario ke 3_2 ini menggunakan jumlah training 32 (Januari 2014 - Agustus 2016) dan testing 15 (September 2016 – November 2017) dengan epoch yang digunakan pada ANN sebesar 500 epoch. Epoch adalah iterasi untuk pengulangan pembaharuan bobot. Nilai epoch ditentukan dengan *trial* dan *error* sampai nilai tiap-tiap bobot sama dengan nilai bobot dari iterasi sebelumnya. Dari pengujian ini maka dapat disimpulkan bahwa prediksi harga bawang merah sangat dipengaruhi oleh atribut penguapan (PU), dan curah hujan (CH). Setelah didapatkan hasil yang terbaik dari hibridisasi CART-ANN, kemudian dilakukan pengujian terhadap CART dan ANN secara konvensional dengan menggunakan semua atribut untuk dibandingkan dengan hibridisasi CART-ANN yang menggunakan atribut terpenting saja. Hal ini dilakukan untuk membuktikan bahwa hasil hibridisasi CART-ANN dengan menggunakan atribut terpenting saja dapat memberikan kinerja sistem yang lebih baik. Hasil *precision*, *recall*, dan akurasi dari hibridisasi CART-ANN terbukti dapat memberikan performansi sistem yang lebih baik daripada algoritma CART atau ANN secara konvensional sebagaimana dapat dilihat pada gambar 3 di bawah ini.



Gambar 3. Perbandingan Hasil Pengujian dari hibridisasi CART-ANN, CART dan ANN konvensional

Jumlah data *training* dan *testing* yang digunakan pada pengujian hibridisasi CART-ANN, CART dan ANN konvensional sama, yaitu 32 data *training* (Januari 2014 - Agustus 2016) dan 15 data *testing* (September 2016 – November 2017) namun atribut yang digunakan berbeda, yaitu untuk hibridisasi CART-ANN hanya menggunakan atribut penting dari proses CART kemudian dilanjutkan ke proses inputan ANN, sedangkan untuk metode CART dan ANN secara konvensional menggunakan semua atribut. Hal ini membuktikan bahwa pemilihan atribut penting ini sangat berpengaruh untuk menjadikan sistem lebih baik daripada menggunakan semua atribut. Selain itu dengan ditambah inputan dari kelas prediksi dari CART sebagai inputan tambahan di ANN dapat meningkatkan dimensi ruang fitur dan juga meningkatkan pemisahan kelas karena semakin banyak informasi.

4.2. Analisis Pengujian

Pada penelitian ini dilakukan tiga skenario, yaitu skenario pembagian data *training* dan *testing*, skenario untuk menentukan kedalaman maksimum *tree*, dan ketiga adalah skenario perbandingan antara metode hibridisasi CART-ANN yang hanya menggunakan atribut terpenting saja pada masukannya dengan metode ANN atau CART secara konvensional dengan menggunakan semua atribut pada masukannya. Setelah dilakukan pengujian, dapat dianalisa bahwasanya skenario partisi data ketiga dengan nilai kedalaman maksimum *tree* dua memberikan hasil hibridisasi CART-ANN yang lebih baik dari pada skenario lainnya. Sebagaimana dapat dilihat pada tabel 7 dengan atribut terpenting yang dihasilkan yaitu atribut penguapan (PU), curah hujan (CH) dan harga pasar aktual (HPA). Perbandingan jumlah *training* dan *testing* yang digunakan yaitu 32 data *training* (Januari 2014 - Agustus 2016) dan 15 data *testing* (September 2016 – November 2017). Penentuan jumlah *training* dan *testing* ini dilakukan dengan *trial* dan *error*, karena jika *training* terlalu sedikit maka sistem akan sedikit untuk belajar, namun jika data *training* terlalu banyak maka sistem akan sedikit dalam pengujian data. Penentuan nilai kedalaman maksimum pada CART juga harus dilakukan dengan *trial* dan *error* karena jika nilai kedalaman pohon maksimum nya terlalu tinggi maka akan menghasilkan model yang kompleksitas nya tinggi, sedangkan jika terlalu rendah maka akan menyebabkan penghapusan beberapa parameter input. Pada skenario perbandingan hibridisasi CART-ANN dengan algoritma CART atau ANN secara konvensional dapat dilihat pada gambar 3 bahwasanya untuk evaluasi performansi dengan menggunakan *confusion matrix* berupa *precision*, *recall*, dan akurasi didapatkan bahwa hibridisasi CART-ANN memberikan kinerja yang lebih baik yang kemudian disusul oleh ANN dan terakhir yaitu CART. Sehingga dapat diambil kesimpulan bahwa pemilihan atribut penting sebagai inputan terbukti lebih efektif dari pada pemakaian semua atribut.

5. Kesimpulan

Berdasarkan beberapa skenario pengujian dalam penelitian ini, metode hibridisasi CART-ANN terbukti memberikan hasil yang lebih baik dari pada CART atau ANN secara konvensional dengan nilai *precision*, *recall*, dan akurasi berturut-turut sebesar 90.91%, 100.00%, 93.33%. skenario partisi data ketiga dengan nilai kedalaman maksimum *tree* dua memberikan hasil hibridisasi CART-ANN yang lebih baik dari pada skenario lainnya dengan jumlah *training* dan *testing* yang digunakan yaitu 32 data *training* (Januari 2014 - Agustus 2016) dan 15 data *testing* (September 2016 – November 2017). Pada proses hibridisasi CART-ANN ini diperoleh atribut penting untuk prediksi harga terhadap cuaca yaitu atribut penguapan (PU), curah hujan (CH) dan harga pasar aktual (HPA).

Beberapa saran yang bisa dilakukan untuk penelitian lebih lanjut diantaranya adalah dengan melakukan skenario yang lebih banyak lagi pada partisi data *training* dan *testing*. Selain itu, dapat juga dicoba menggunakan faktor yang mempengaruhi pasokan bawang merah lainnya seperti faktor pemilihan varietas benih dan evaluasi kesesuaian lahan.

Daftar Pustaka

- [1]. Sari, R. M. (2016). SUPPLY ANALYSIS OF RICE, RED CHILI AND SHALLOT IN INDONESIA. *JURNAL AGRIBISNIS TERPADU*, 9(1).
- [2]. Zeppa, V. S. (2016). Institutional Innovation in Farmer-Trader Exchange Practices in the Indonesian Shallot Market.
- [3]. Sukasih, E., & Musadad, D. (2018, January). Physico-chemical characteristics of shallot New-Superior Varieties (NSV) from Indonesia. In *IOP Conference Series: Earth and Environmental Science* (Vol. 102, No. 1, p. 012037). IOP Publishing.
- [4]. Haryanti, D. M., Kusumawardhani, M., & Anshar, K. (2018). Indonesia: Inclusive Business in Indonesia—Improving Supply Chain Efficiency through Inclusive Business.
- [5]. Sari, Y. R., Manullang, N., Farliani, N., Oktarina, M., Rahman, F., Jaeroni, A., ... & Bintara, H. (2015). *Strategic Study On Cluster Strengthening To Support The Supply Of Volatile Food Commodities* (No. WP/8/2014).
- [6]. Prabawati, N. F., Pujawan, I. N., & Widodo, E. (2018, April). Modeling of shallot supply decisions: the case of Indonesia. In *IOP Conference Series: Materials Science and Engineering* (Vol. 337, No. 1, p. 012016). IOP Publishing.
- [7]. Nhita, F., Saepudin, D., & Wisesty, U. N. (2018). Planting Date Recommendation for Chili and Tomato Based on Economic Value Prediction of Agricultural Commodities. *The Open Agriculture Journal*, 12(1).
- [8]. Nuvaishyah, P., Nhita, F., & Saepudin, D. (2019). Price Prediction of Chili Commodities in Bandung Regency Using Bayesian Network. *International Journal on Information and Communication Technology (IJICT)*, 4(2), 19-32.
- [9]. Asma Hasifa Nurcahyono, Fhira Nhita, Deni Saepudin, and Annisa Aditsania. "Price Prediction of Chili in Bandung Regency Using Support Vector Machine (SVM) Optimized with an Adaptive Neuro-Fuzzy Interference System (ANFIS) (Proses Publish).
- [10]. Chakraborty, T., Chattopadhyay, S., & Chakraborty, A. K. (2018). A novel hybridization of classification trees and artificial neural networks for selection of students in a business school. *Opsearch*, 55(2), 434-446.
- [11]. Salimi, A., Faradonbeh, R. S., Monjezi, M., & Moormann, C. (2018). TBM performance estimation using a classification and regression tree (CART) technique. *Bulletin of Engineering Geology and the Environment*, 77(1), 429-440.
- [12]. Kim, K. (2016). A hybrid classification algorithm by subspace partitioning through semi-supervised decision tree. *Pattern Recognition*, 60, 157-163.
- [13]. Li, F., Zhang, X., Du, C., & Huang, L. (2015, November). A hybrid NRS-CART algorithm and its application on coal mine floor water-inrush prediction. In *TENCON 2015-2015 IEEE Region 10 Conference* (pp. 1-4). IEEE.
- [14]. Suyatno, J. A., Nhita, F., & Rohmawati, A. A. (2018, May). Rainfall Forecasting in Bandung Regency Using C4. 5 Algorithm. In *2018 6th International Conference on Information and Communication Technology (ICICT)* (pp. 324-328). IEEE.
- [15]. Suyanto, D. (2017). Data Mining untuk klasifikasi dan klusterisasi data. *Bandung: Informatika Bandung*.
- [16]. Puspitaningrum, D. (2006). Pengantar Jaringan Syaraf Tiruan. *Yogyakarta: Andi*.
- [17]. Hermawan, A. (2006). Jaringan Saraf Tiruan Teori dan Aplikasi. *Yogyakarta: Andi*.
- [18]. Sri, K., & Sri, H. (2010). Neuro-Fuzzy Integrasi Sistem Fuzzy dan Jaringan Syaraf. *Yogyakarta: Penerbit Graha Ilmu*.
- [19]. García, S., Luengo, J., & Herrera, F. (2015). *Data preprocessing in data mining* (pp. 59-139). New York: Springer.
- [20]. Nhita, F., Annisa, S., & Kinasih, S. (2015, May). Comparative study of grammatical evolution and adaptive neuro-fuzzy inference system on rainfall forecasting in Bandung. In *2015 3rd International Conference on Information and Communication Technology (ICICT)* (pp. 6-10). IEEE.
- [21]. Han, J., Kamber, M., & Pei, J. (2011). Data mining concepts and techniques third edition. *The Morgan Kaufmann Series in Data Management Systems*, 83-124.