

## **K-NEAREST NEIGHBOR DENGAN FEATURE SELECTION MENGUNAKAN BACKWARD ELIMINATION UNTUK PREDIKSI HARGA KOMODITI KOPI ARABIKA**

**Andi Bode**

andibode22@gmail.com  
Universitas Ichsan Gorontalo

### **Abstrak**

Kopi arabika tergolong salah satu komoditas unggulan didalam subsektor perkebunan di Indonesia karena memiliki peluang pasar yang baik di dalam negeri maupun luar negeri. Dalam penelitian ini akan dilakukan peramalan harga komoditi kopi arabika. Metode *Time series* adalah metode yang digunakan untuk peramalan dimasa lalu dan mengetahui nilai di masa yang akan datang. Seleksi fitur digunakan sebagai tujuan untuk memilih variabel-variabel yang signifikan dalam melakukan prediksi harga komoditi kopi arabika menggunakan *K-Nearest Neighbor* (KNN) dengan *Backward Elimination* (BE). Hasil eksperimen penelitian ini menunjukkan dimana algoritma KNN dengan *Backward Elimination* dapat memperkecil nilai *error*, dibandingkan dengan KNN tanpa seleksi fitur dan BPNN, BPNN dengan *Backward Elimination*.

**Kata kunci:** Prediksi, *K-Nearest Neighbor*, *Backward Elimination*, *Time Series*

### **1. Pendahuluan**

Sebagian besar wilayah Indonesia adalah wilayah lahan pertanian, sehingga Negara Indonesia terkenal dengan Negara agraris. Potensi pendapatan sumber daya terbesar didapatkan dari sektor pertanian. Perkebunan, sebagai salah satu subsektor penting dalam sektor pertanian, mempunyai kontribusi yang signifikan terhadap perekonomian Indonesia [1]. Kopi arabika tergolong salah satu komoditas unggulan didalam subsektor perkebunan di Indonesia karena memiliki peluang pasar yang baik di dalam negeri maupun luar negeri. Produksi kopi Indonesia telah mencapai 600 ribu ton pertahunnya dan lebih dari 80 persen berasal dari perkebunan rakyat [2]. Fluktuasi harga kopi yang terjadi tidak terlepas dari perkembangan produksi kopi dunia karena adanya faktor penyebab. Pesatnya perkembangan produksi kopi dunia telah menyebabkan terjadinya pasokan kopi dunia sehingga mengakibatkan persaingan antara negara produsen semakin ketat dan pada akhirnya harga cenderung tertekan [3]. Metode *time series* sebagai metode peramalan, data *time series* pada umumnya untuk mengetahui bentuk pola variasi dimasa lalu dan mengetahui nilai di masa yang akan datang. Dengan melakukan proses determinasi metode *time series* memiliki kemampuan prediksi yang tinggi, dengan meninjau nilai di masa lalu maka nilai yang akan datang dapat diketahui. Signal yang kuat pada komponen determinan maka nilai yang akan datang dapat diketahui [4].

Metode K-NN adalah merupakan salah satu pendekatan yang digunakan dalam pengklasifikasian dengan secara mudah dan efisien. Konsep mendasar dari algoritma KNN yaitu mencari jarak paling terdekat diantara data yang terevaluasi dengan sejumlah *K* tetangga (neighbor) paling dekat dalam data uji [5]. KNN bekerja dengan cara membandingkan data uji dan data training/template. KNN mencari pola data template yang paling mendekati dengan data uji. KNN memiliki kelebihan kepada training data set yang memiliki banyak noisy serta efektif terhadap jumlah data training tinggi/besar. Namun kekurangan dari KNN masih perlu penentuan nilai *K* dan untuk pemilihan atribut terbaik [6]. *Feature Selection* adalah proses pemilihan suatu subset dari fitur asli dengan menghilangkan fitur yang tidak relevan. Pemilihan fitur akan meningkatkan efisiensi. Tujuan utama dari seleksi fitur adalah fokus pada mencari sebuah data yang relevan. Fitur yang tidak relevan dan fitur berlebihan berpengaruh terhadap hasil sehingga pemilihan fitur harus bisa mengidentifikasi data tersebut [7]. *Backward Elimination* untuk menghilangkan atribut atau variabel yang tidak relevan, apabila ada variabel dianggap tidak berpengaruh atau tidak signifikan dalam model maka akan dihapus dari model [8]. Metode *backward elimination* menghasilkan performa kinerja lebih baik ketika dibandingkan dengan cara statistik dalam menyeleksi fitur. Kinerja terbaik dapat di peroleh dari sensitivitas, spesifisitas dan keakuratan [9]. Pada penelitian ini bertujuan untuk menghasilkan model prediksi

harga komoditi kopi arabika dengan tingkat *error* lebih rendah dengan menerapkan *Feature Selection* menggunakan *Backward Elimination* pada *K-Nearest Neighbor* dan dibandingkan dengan BPNN.

## 2. Metode

### 2.1. Data Maining

Teknik yang dibuat didalam data mining yaitu salah satunya bagaimana menelusuri data yang ada untuk membentuk sebuah model, kemudian menggunakan model tersebut bisa dapat mengenali pola data yang lain yang tidak berada dalam basis data yang sudah tersimpan. Teknik seperti ini juga dapat dimanfaatkan untuk kebutuhan prediksi. Dalam data mining, pengategorisasian data juga bisa dilakukan. Bertujuan agar kita dapat mengetahui pola universal data-data yang ada [10].

### 2.2. Time Series

Metode *time series* merupakan metode peramalan dengan menggunakan analisa pola hubungan antara variabel yang akan diperkirakan dengan variabel waktu atau analisis *time series*. Klasifikasi dari model runtun waktu dapat dikelompokkan menjadi dua. Pertama model univariat: hanya mengamati satu variabel/individu runtun waktu. Kedua model multivariat (multivariabel) : lebih dari satu variabel/individu runtun waktu. *Time series* multivariate memodelkan peubah-ubah yang berkolerasi dan tertulis dari waktu ke waktu [11].

**Tabel 1.** Pola Data Time Series

Pola	Input lag	Output/target
1	X <sub>1</sub> , X <sub>2</sub> , X <sub>3</sub> , X <sub>4</sub> , ..., X <sub>p</sub>	X <sub>p+1</sub>
2	X <sub>2</sub> , X <sub>3</sub> , X <sub>4</sub> , X <sub>5</sub> , ..., X <sub>p+1</sub>	X <sub>p+2</sub>
3	X <sub>3</sub> , X <sub>4</sub> , X <sub>5</sub> , X <sub>6</sub> , ..., X <sub>p+2</sub>	X <sub>p+3</sub>
...	...	...
m-p	X <sub>m-p</sub> , X <sub>m-p+1</sub> , X <sub>m-p+2</sub> , ..., X <sub>m-1</sub>	X <sub>m</sub>

### 2.3. Seleksi Fitur

*Backward Elimination* merupakan suatu metode yang memiliki fungsi untuk mengoptimalkan kinerja suatu model dengan sistem kerja pemilihan mundur. Pemilihan variabel dilakukan dengan cara pemilihan kedepan yakni menguji semua variabel kemudian menghapus variabel-variabel yang dianggap tidak signifikan. Variabel yang diproses satu per satu, jika variabel dianggap tidak berpengaruh atau tidak signifikan dalam model maka akan dihapus dari model.

### 2.4. K-Nearest Neighbor

Algoritma *K-Nearest Neighbor* (KNN) adalah salah satu metode yang menerapkan algoritma *supervised* yang mana algoritma tersebut terbagi atas dua jenis yaitu *supervised learning* dengan *unsupervised learning*. Algoritma *supervised learning* bertujuan untuk mendapatkan pola baru sedangkan *unsupervised learning* untuk mendapatkan pola dalam sebuah data [12]. Ketepatan algoritma KNN ditentukan oleh ada dan tidak adanya data yang tidak relevan, atau jika bobot fitur tersebut setara dengan relevansinya terhadap klasifikasi [13]. Algoritma KNN adalah salah satu metode yang digunakan untuk analisis klasifikasi, namun beberapa dekade terakhir metode KNN juga digunakan untuk prediksi [14]. Mencari jarak terdekat antara data yang akan dievaluasi dengan *k* tetangga (*neighbor*) terdekatnya dalam data pelatihan. Ruang ini dibagi menjadi kelompok-kelompok berdasarkan klasifikasi data training. Sebuah titik pada tempat ini ditandai dengan kelas *c*, dimana jika kelas *c* merupakan klasifikasi yang sangat banyak ditemui pada *k* buah tetangga terdekat titik tersebut. Untuk sistem kerja KNN data training di proyeksikan ke tempat berdimensi banyak, yang mana dari masing-masing dimensi mempersentasikan fitur dari data [15]. Model persamaan algoritma KNN dibawah ini:

$$D = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} \quad (1)$$

Keterangan:

$x$  = sampel data  
 $y$  = data uji  
 $D$  = jarak

$$fknn(x') = \frac{1}{K} \sum_{i \in N_{k(x')}} y_i \quad (2)$$

$x'$  = Perkiraan atau estimasi  
 $K$  = Jumlah tetangga terdekat  
 $N_k(x')$  = Tetangga terdekat  
 $y_i$  = Output tetangga terdekat

## 2.5. Backpropagation Neural Network

*Neural Network* sebagai salah satu metode pembelajaran yang mana diinspirasi dari sebuah jaringan sistem pembelajaran biologis yang terdiri dari jaringan saraf (neuron) yang terhubung antara satu dengan yang lainnya. Struktur *neural network* yang dipakai yaitu *backpropagation* dimana merupakan sebuah metode sistematis untuk pelatihan *multilayer*. Metode ini memiliki dasar sistematis yang kuat, objektif, dan algoritma ini mendapatkan bentuk persamaan dan nilai koefisien dalam formula dengan meminimalkan jumlah kuadrat galat error [16]. *Backpropagation* banyak digunakan untuk menyelesaikan permasalahan, salah satunya yaitu prediksi. Cara pelatihan dalam metode ini sama dengan *Perceptron* adalah sejumlah data latih sebagai pola masukan dikirimkan ke jaringan kemudian dihitung pola keluaran dan jika terdapat *error* maka bobot dalam jaringan tersebut akan diperbaharui guna mengurangi *error* tersebut. *Error* tersebut merupakan perbedaan antara target keluaran yang diinginkan dengan nilai keluaran [17].

## 2.6. K-fold Cross Validation

*K-Fold Cross Validation* termasuk juga suatu teknik validasi untuk membagi data ke dalam  $k$  bagian dan selanjutnya masing-masing bagian tersebut masuk dalam proses klasifikasi. Dimana dengan memakai *K-Fold Cross Validation* akan dilakukan testing sebanyak  $k$ . Tiap testing tersebut menggunakan satu data testing dan  $k-1$  bagian akan menjadi data training, kemudian data testing itu akan ditukarkan dengan satu buah data training sehingga untuk tiap testing akan diperoleh data testing yang berbeda-beda [18].

## 2.7. Root Mean Square Error (RMSE)

*Root Mean Square Error* (RMSE) merupakan sebuah ukuran yang sering digunakan dari perbedaan antara nilai-nilai diprediksi oleh model atau estimator dan nilai-nilai benar-benar diamati. Perbedaan-perbedaan individual disebut residual ketika perhitungan tersebut dilakukan untuk sampel data yang digunakan buat estimasi, dan disebut juga sebuah kesalahan dalam melakukan prediksi, ketika dihitung *out-of-sample*. RMSE merupakan suatu cara yang sering digunakan didalam pengevaluasian hasil peramalan yaitu dengan menggunakan RMSE, error yang ada menjelaskan seberapa besar perbedaan hasil estimasi dengan hasil yang akan di estimasi [19]. Nilai RMSE rendah membuktikan bahwa variasi nilai yang dihasilkan oleh suatu model prakiraan mendekati variasi nilai observasinya. Menurut Raharja, A [20]. berikut persamaan *root means square* (RMSE).

$$RMSE = \sqrt{\frac{\sum (y_t - \hat{y}_t)^2}{n}} \quad (3)$$

Dimana:

**RMSE** = Root Mean Square Error  
 $n$  = Jumlah Sampel  
 $y_t$  = Nilai Aktual Indeks  
 $\hat{y}_t$  = Nilai Prediksi Indeks

### 3. Hasil dan Pembahasan

Pada penelitian ini memiliki tujuan untuk menerapkan algoritma KNN menggunakan *Backward Elimination* pada data Komoditi yang bertujuan untuk mencari model yang terbaik dalam melakukan prediksi *time series* harga komoditi. Sumber data pada penelitian ini di ambil dari Badan Pengawas Perdagangan Berjangka Komoditi (BAPPEBTI). Data set tersebut berupa data set harian harga kopi arabika tahun 2005-2015 dengan jumlah record data 2598, data set harian kemudian diubah menjadi data bulanan sehingga diperoleh 132 record. Dimana data set tersebut tidak terjadi transaksi di hari sabtu dan minggu dan tanggal merah kalender. Data set tersebut terdiri dari variabel harga/kg.

Proses pengolahan dataset, data yang akan diolah berupa data univariat *time series* berupa numerik data harian. Data *univariat* menjadi *multivariate* yang terdiri dari 1 variabel periode atau independen sampai 7 variabel independen dengan masing-masing 1 variabel dependen. Sebelum mengubah data menjadi beberapa variabel, terlebih dahulu susunan datanya diubah dari *ascending* menjadi *descending* kemudian data akan dinormalisasi, fungsi normalisasi pada ilmu database atau basis data yaitu untuk menghindari terjadinya berbagai anomaly data dan tidak konsistensinya data. Tujuan dari normalisasi untuk memperoleh data yang memiliki ukuran lebih kecil untuk mewakili data asli dengan tidak kehilangan karakteristik dirinya [21].

Rumus Normalisasi Data :

$$\text{Normalisasi} = \frac{(X - \min)}{(max - \min)} \quad (4)$$

Dimana:

$x$  = Data

$Min$  = Data Minimum

$Max$  = Data Maksimum

Denormalisasi dataset adalah mengembalikan ukuran dataset yang sudah ternormalisasi untuk mengembalikan data asli. Denormalisasi diterapkan pada hasil eksperimen data testing yang berupa prediksi harga komoditi kopi arabika.

Rumus Denormalisasi Data:

$$\text{Denormalisasi} = Y (max - min) + min \quad (5)$$

Dimana:

$Y$  = Hasil Keluaran Dari Pelatihan

$Min$  = Data Minimum

$Max$  = Data Maksimum

Proses seleksi variabel, digunakan *Backward Elimination*. Fungsi *Backward Elimination* untuk menghilangkan atribut-atribut yang tidak relevan. Data yang ditraining dilakukan secara bertahap dimulai dari 1 variabel periode sampai pada tingkat atau jumlah variabel periode yang menghasilkan performa, nilai akurasi paling baik atau *error* terkecil untuk membandingkan data mana yang menghasilkan prediksi paling baik. Tahapan ini yang digunakan dalam pengujian yaitu *K-Nearest Neighbort*. Algoritma KNN sebagai metode yang menggunakan algoritma *supervised* dimana algoritma tersebut terbagi atas dua jenis yaitu *supervised learning* dan *unsupervised learning*. Tujuan dari *supervised learning* untuk menghasilkan pola baru sedangkan tujuan dari *unsupervised learning* untuk menghasilkan pola dalam sebuah data. Pada pengujiannya untuk setiap data yang diuji agar menemukan hasil yang maksimal dalam setiap pengujian data, nilai  $k$  dari algoritma ini dilakukan perubahan sampai menemukan nilai yang paling baik. Sehingga diketahui data yang mana dan nilai  $k$  berapa yang menghasilkan nilai yang paling baik. Selain algoritma KNN peneliti juga menggunakan algoritma BPNN sebagai pembanding yang paling baik dalam memprediksi harga kopi arabika.

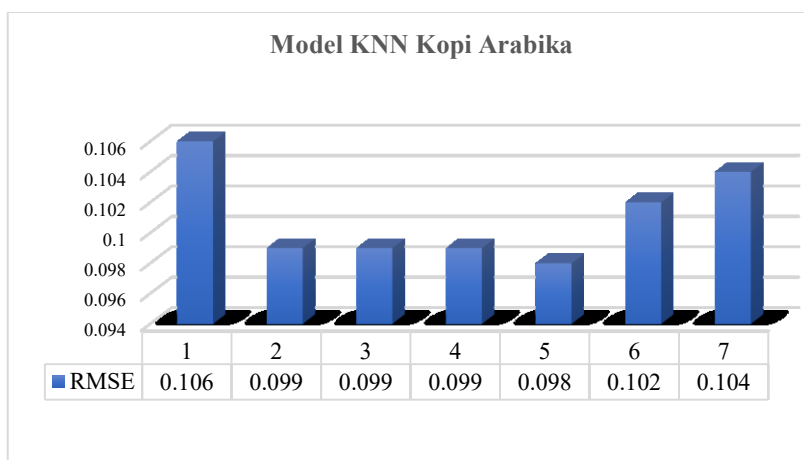
Pada penelitian ini eksperimen dan pengujian dilakukan beberapa tahap, tahap pertama mengubah data harian menjadi data bulanan, kemudian data dinormalisasi. Tahap kedua mengolah data univariat ke multivariat. Tahap ketiga menentukan parameter KNN, KNN menggunakan *Backward Elimination*. Tahap keempat menentukan parameter BPNN, dalam menentukan parameter BPNN hampir sama dengan tahap sebelumnya yaitu dilakukan beberapa pengujian penentuan jumlah

variabel input atau variabel periode, tetapi tidak melakukan penentuan nilai  $k$ . Pengujian tersebut dilakukan untuk mendapatkan model yang terbaik dengan melihat nilai RMSE terkecil. Tahap kelima mengaplikasikan model yang terbaik antara model yang dihasilkan algoritma KNN dan BPNN menggunakan *Backward Elimination* terhadap data testing untuk melakukan prediksi. Untuk proses eksperimen, digunakan beberapa sampel data set untuk melakukan uji coba, sebelum melakukan uji coba maka dilanjutkan dengan penentuan parameter, nilai parameter yang dipakai dalam melakukan proses uji coba yaitu 1, 3, 5, 7, 9, 11, 13, 15, *number of validation* 10, nilai ketetanggaan atau  $K$  1, lalu dilakukan proses testing untuk mendapatkan *root mean square error*.

**Tabel 2.** Nilai RMSE Kopi Arabika dengan KNN

Variabel Periode	Of Validation	K	RMSE
1	10	15	0.106
2	10	15	0.099
3	10	13	0.099
4	10	13	0.099
5	10	11	0.098
6	10	13	0.102
7	10	9	0.104

Pada Tabel 2. Merupakan tabel rangkuman hasil eksperimen kopi arabika dari penentuan model dengan menggunakan algoritma KNN yang dimulai dari 1 sampai 7 variabel periode dengan menggunakan *cross validation* 10 dan nilai  $K$  1, 3, 5, 7, 9, 11, 13, 15 yang dipilih berdasarkan nilai RMSE terkecil. Maka didapatkan model terbaik yang ditunjukkan berdasarkan nilai error terkecil 0.098 dengan jumlah variabel periode 5 jumlah validasi 10 dan nilai  $k$  yaitu 11.



**Gambar 1.** Grafik Model KNN Kopi Arabika

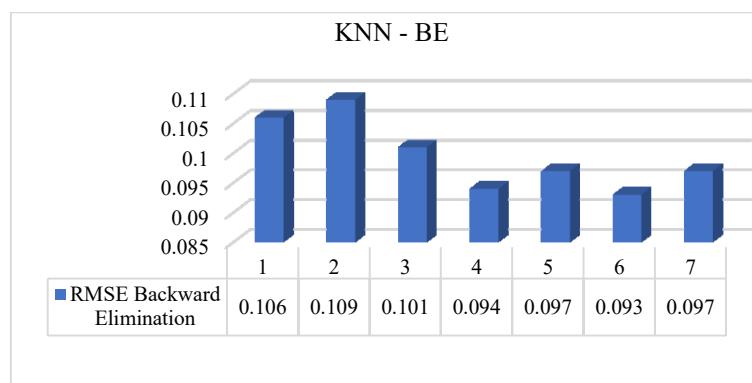
Pada gambar 1. Merupakan grafik model terbaik untuk data kopi arabika, menunjukkan bahwa model yang paling baik terdapat pada variabel periode 5, nilai validasi 10 dan  $K$  11, untuk data kopi arabika dengan nilai rmse-nya sebesar 0.098.

**Tabel 3.** Nilai RMSE Kopi Arabika dengan KNN menggunakan *Backward Elimination*

Variabel Periode	Of Validation	K	RMSE Backward Elimination
1	10	15	0.106
2	10	5	0.109

3	10	13	0.101
4	10	13	0.094
5	10	13	0.097
6	10	15	0.093
7	10	11	0.097

Pada Tabel 3. Tersebut merupakan hasil rangkuman eskperimen dari penentuan model dengan algoritma KNN menggunakan *Backward Elimination* yang dimulai dari 1 sampai 7 variabel periode dengan *cross validation* 10 dan nilai  $K$  1, 3, 5, 7, 9, 11, 13, 15 yang dipilih berdasarkan nilai RMSE terkecil. Maka didapatkan model terbaik yang ditunjukkan berdasarkan nilai error terkecil yaitu 0.093 dengan jumlah variabel periode 6, jumlah validasi 10 dan nilai  $k$  yaitu 15, adapun variabel yang signifikan yaitu  $xt-1$ ,  $xt-2$  dan  $xt-3$ .

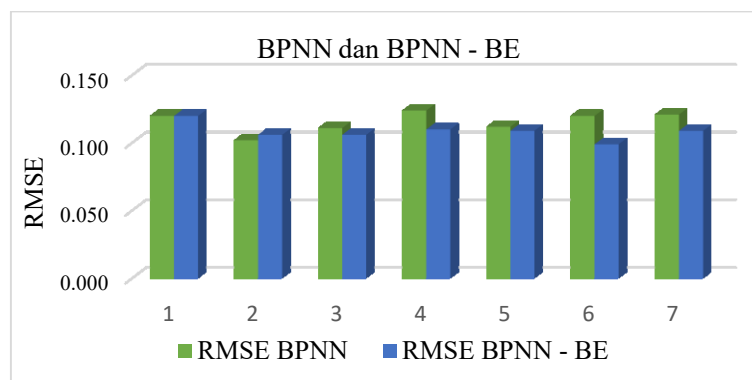


**Gambar 2.** Grafik Model KNN – *Backward Elimination*

Berdasarkan gambar 2. Sebagai grafik model terbaik kopi arabika, memperlihatkan bahwa model yang paling baik terdapat pada variabel periode 6, dengan menggunakan *Backward Elimination* yaitu nilai rmse-nya sebesar 0.093.

### 3.1. Perbandingan Model

Pada penelitian ini juga melakukan eksperimen menggunakan model BPNN, BPNN dan *Forward* Penelitian ini juga melakukan eksperimen menggunakan algoritma *Backpropagation Neural Network*, *Backpropagation Neural Network* menggunakan *Backward Elimination*. Hasil perbandingan RMSE model *Backpropagation Neural Network*, *Backpropagation Neural Network* menggunakan *Backward Elimination* diperlihatkan pada gambar 4.



**Gambar 3.** Grafik Model BPNN dan BPNN – BE Kopi Arabika

Berdasarkan gambar 3. Sebagai grafik model terbaik kopi arabika, memperlihatkan bahwa model yang paling baik terdapat pada variabel periode 6, dengan algoritma BPNN menggunakan *Backward Elimination* yaitu nilai rmse-nya sebesar 0.099, pada parameter *hidden layer 1 size 10 training cycle 1100 learning rate 0.1 momentum 0.0*. Hasil perbandingan prediksi RMSE algoritma KNN dan BPNN diperlihatkan pada tabel 4.

**Tabel 4.** Hasil Perbandingan Algoritma KNN, KNN BE, BPNN, BPNN BE

Kopi Arabika		
Algoritma	Variabel Periode	RMSE
K-NN	5	0.098
K-NN – Backward Elimination	6	0.093
BPNN	2	0.102
BPNN – Backward Elimination	6	0.099

Tabel 4. Sudah dapat dilihat model yang akan digunakan untuk memprediksi harga komoditi kopi arabika yaitu model pada algoritma KNN menggunakan *Backward Elimination*. Dari tabel diatas maka model yang paling baik dapat dilihat dari tingkat *error* terkecil. Dengan penambahan suatu fitur seleksi dapat meningkatkan performa suatu algoritma.

#### 4. Kesimpulan dan Saran

Peramalan harga komoditi kopi arabika dengan menggunakan algoritma KNN dengan *Feature Selection* menggunakan *Backward Elimination* berhasil dilakukan. Berdasarkan hasil eksperimen yang telah dilakukan, menunjukkan bahwa seleksi fitur yakni *Backward Elimination* meningkatkan performa yang lebih baik. Kisaran nilai prediksi harga kopi arabika menghampiri dengan harga komoditi kopi arabika sebenarnya, dengan nilai hasil perbandingan dibulan januari 91.6% dan february 95.3%. Algoritma KNN menggunakan *Backward Elimination* lebih baik dibandingkan algoritma KNN dan BPNN, BPNN menggunakan *Backward Elimination*.

Disarankan untuk penulis selanjutnya untuk menggunakan algoritma lain dengan harapan menghasilkan tingkat *error* lebih kecil.

#### Daftar Pustaka

- [1] P. Septianingtyas Ari. 2015. Faktor-faktor yang mempengaruhi harga kopi arabika dan integrasi pasar kopi arabika dunia terhadap kopi arabika di Indonesia, universitas jember.
- [2] Ilham, Supriana T, Salimah. 2013. Analisis faktor-faktor yang mempengaruhi permintaan komoditi kopi di Sumatera utara, Fakultas Pertanian USU.
- [3] Kustiari Reni. 2007. Perkembangan pasar kopi dunia dan implikasinya bagi Indonesia, Bogor : Forum penelitian argo ekonomi. Volume 25 No. 1, 43 – 55.
- [4] Sularno Aris. 2012. Prediksi nilai saham menggunakan *pemrograman genetic* dan *pemrograman ekspresi gen*, universitas gunadarma.
- [5] Sonatha Y. 2013. Churn prediction pelanggan menggunakan Crisp-DM Studi Kasus Pelanggan TelkomFlexi Bandung, *Jurnal Elektron*. Vol. 5. No. 1, UNAND Limau Manis, Padang.
- [6] Bahri R. sofian. 2012. Perbandingan algoritma template matching dan feature extraction pada optical character recognition. *Jurnal Computer Dan Informatika*. Universitas Computer Indonesia, Edisi. I volume. 1.
- [7] Poonguzhali. E, Vijayalakshmi. M, Shamily. K, Priyadarshini. V. 2014. Hybrid feature selection algorithm for high dimensional database, Pondicherry university india, *International Journal Of Engineering Trends And Technology* (IJETT) – Volume 8 Number 9.
- [8] L. Ladha, T. Deepa. 2011. Feature Selection Methods And Algorithms, *International Journal on Computer Science and Engineering* (IJCSE), ISSN : 0975-3397 vol. 3 No. 5.

- [9] Ali Narin, Yalcin Isler dan Mahmut Ozer. 2014. Investigating the performance improvement of HRV indices in CHF using feature selection methods based on backward elimination and statistical significance, computer in biology and medicine.
- [10] Prasetyo Eko. 2012. Data mining konsep dan aplikasi menggunakan matlab, *ANDI*. Yogyakarta.
- [11] Purwanto, C. Eswara, and R. Logeswara. 2011. Improved Adaptive *Neuro-fuzzy* Inference System for HIV/AIDS *Time series* Prediction. In: *Informatics Engineering and Information Science*, 253, Springer-Verlag Berlin.
- [12] Krisandi Nobertus, Helmi, Prihandono Bayu. 2013. Algoritma *k-nearest neighbor* dalam klasifikasi data hasil produksi kelapa sawit pada Pt.Minamas kecamatan parindu, *Bimaster* : volume 02, No.1.
- [13] Nugroho, R.S, Wijana Katon. 2015. Program bantu prediksi penjualan barang menggunakan metode KNN studi kasus: U.D. Anang, *Jurnal Eksis*. Vol 08 No 02 November.
- [14] Alkhatib K, Najadat H, Hmeidi I, Shatnawi MKA. 2013. Stock Price Prediction Using *k*-Nearest Neighbor (*k*NN) Algorithm. *International Journal of Business, Humanities and Technology*. 3(3):32-44.
- [15] R. N. Whidhiasih, N. A. Wahanani dan Supriyanto. 2013. Klasifikasi buah belimbing berdasarkan citra red-green-blue menggunakan knn dan Ida, *Jurnal Penelitian Ilmu Komputer, System Embedded & Logic* 1(1) : 29-35.
- [16] Novi, Y. 2011. Penerapan Metode Neural Network dengan Struktur Backpropagation untuk Prediksi Stok Obat di Apotik, Seminar nasional aplikasi teknologi informasi.
- [17] Raharjo, J. S. Dwi. 2013. Model Artificial Neural Network berbasis Particle Swarm Optimization untuk prediksi laju inflasi, *Jurnal Sistem Komputer* - Vol. 3, No.1, ISSN: 2087-4685, e-ISSN: 2252-3456.
- [18] Zega S. Artaty. 2014. Penggunaan pohon keputusan untuk klasifikasi tingkat kualitas mahasiswa berdasarkan jalur masuk kuliah, Politeknik Negeri Batam, Yogyakarta. ISSN: 1907 – 5022.
- [19] Fikri, A. 2013. Penerapan Data Mining Untuk Mengetahui Tingkat Kekuatan Beton yang Dihasilkan dengan Metode Estimasi menggunakan Linear Regression, Universitas Dian Nuswantoro, Semarang.
- [20] Alda Raharja, Wiwik Anggraeni, S.Si, M.Kom, Retno Aulia Vinarti, S.Kom. 2010. Menerapkan Metode *Exponential Smoothing* Untuk Peramalan Penggunaan Waktu Telpon di PT. Telkom Divre 3, *SISFO-Jurnal Sistem Informasi*.
- [21] Indrabayu, Harun, N, Pallu, M.S, Achmad, A, dan L, F.I. 2012. Prediksi Curah Hujan dengan Jaringan Saraf Tiruan, Universitas Hasanuddin, ISBN: 978-979-127255-0-6, Vol. 6, Makassar.