

SYSTEMIC: Information System and Informatics Journal

ISSN: 2460-8092, 2548-6551 (e)

Vol 6 No 2 – Desember 2020

Klasifikasi Multi Output pada Harga Smartphone Menggunakan Learning Vector Quantization (LVQ) dan Backpropagation (BP)Dinita Rahmalia¹, M. Syaiful Pradana², Teguh Herlambang³^{1,2} Universitas Islam Darul Ulum Lamongan³ Universitas Nahdlatul Ulama Surabayadinitarahmalia@gmail.com¹, syaifulp@unisda.ac.id², teguh@unusa.ac.id³**Kata Kunci**Klasifikasi,
Neural Network,
Learning Vector
Quantization,
Backpropagation,
Pengenalan Pola**Abstrak**

Terdapat banyak smartphone dengan harga yang bermacam-macam yang terjual di pasaran. Harga smartphone dipengaruhi oleh beberapa komponen seperti berat, penyimpanan internal, memory (RAM), kamera belakang, kamera depan dan merek. Terdapat dua metode untuk mengklasifikasikan kelas harga smartphone di pasaran seperti Learning Vector Quantization (LVQ) dan Backpropagation (BP). Dari klasifikasi kelas harga smartphone di pasaran menggunakan LVQ and BP, terdapat perbedaan pada keduanya. LVQ mengklasifikasikan kelas harga smartphone dengan jarak euclidean dari bobot dan data pada iterasinya. BP mengklasifikasikan kelas harga smartphone dengan penurunan gradient dari target dan output pada iterasinya. Pada klasifikasi multi output, satu obyek bisa memiliki lebih dari satu output. Berdasarkan hasil simulasi, BP memberikan akurasi dan error yang lebih baik pada data training dan data testing daripada LVQ.

KeywordsClassification,
Neural Network,
Learning Vector
Quantization,
Backpropagation,
Pattern Recognizing**Abstract**

There are many smartphones with various price sold in market. The price of smartphone is affected by some components such as weight, internal storage, memory (RAM), rear camera, front camera and brands. There are two methods for classifying price class of smartphone in market such as Learning Vector Quantization (LVQ) and Backpropagation (BP). From classifying price class of smartphone in market using LVQ and BP, there are the differences on the both of them. LVQ classifies price range of smartphone by euclidean distance of weight and data on its iteration. BP classifies price range of smartphone by gradient descent of target and output on its iteration. In multi output classification, one object may have multi output. Based on simulation results, BP gives the better accuracy and error rate in training data and testing data than LVQ.

1. Pendahuluan

Pada saat ini perkembangan teknologi informasi dan komunikasi bekerja sangat cepat, luas dan akurat. Salah satu produk yang paling utama dan banyak digunakan adalah smartphone. Smartphone adalah peralatan komunikasi (gadget) yang sangat penting untuk setiap orang sehingga proses komunikasi dapat berjalan sangat cepat. Karena itu, terdapat sangat banyak perusahaan smartphone bersaing memproduksi bermacam-macam smartphone dengan spesifikasi dan harga yang berbeda.

Di pasaran, terdapat banyak smartphone yang dijual dengan harga yang bermacam-macam. Berdasarkan data penjualan smartphone, harga smartphone tergantung pada beberapa komponen

seperti berat, penyimpanan internal, memory (RAM), kamera belakang, kamera depan dan merek, sehingga pada penelitian ini akan dilakukan klasifikasi kelas harga menggunakan metode Neural Network yaitu Learning Vector Quantization (LVQ) dan Backpropagation (BP).

Pada penelitian sebelumnya, aplikasi Neural Network telah diterapkan pada peramalan menggunakan Backpropagation (BP) pada peramalan cuaca [3], peramalan penyebaran penyakit [7], peramalan suhu udara [6] dan Adaptive Neuro Fuzzy Inference System (ANFIS) pada peramalan kelembaban udara [9], peramalan intensitas matahari [11]. Metode peramalan juga diterapkan menggunakan metode Kalman Filter pada peramalan harga saham [10], peramalan

persediaan darah [5]. Selain peramalan, juga terdapat klustering atau pengelompokan menggunakan Fuzzy Clustering Means pada pengelompokan bandara udara [4] dan pengelompokan produksi pertanian [8].

Pada penelitian ini, terdapat dua metode untuk mengklasifikasi harga smartphone di pasaran, yaitu Learning Vector Quantization (LVQ) dan Backpropagation (BP). LVQ dan BP merupakan jenis dari Neural Network yang sering digunakan pada klasifikasi data [1]. Pada proses klasifikasi menggunakan LVQ dan BP, terdapat perbedaan pada kedua metode tersebut. LVQ mengklasifikasi harga smartphone menggunakan jarak euclidean antara bobot (weight) dengan data di setiap iterasi. BP mengklasifikasi harga smartphone menggunakan penurunan gradien (gradient descent) dari target dan output di setiap iterasi. LVQ dan BP bekerja dengan membagi data ke dalam data training dan data testing dengan proporsi yang ditentukan dan mengupdate matriks bobot. Pada LVQ, matriks bobot berperan sebagai vektor referensi sedangkan pada BP, matriks bobot menghubungkan beberapa hidden layer

Pada klasifikasi multi output, satu obyek bisa memiliki lebih dari satu output. Berdasarkan hasil simulasi, BP memberikan accuracy yang lebih baik pada data training maupun data testing daripada LVQ.

2. Metode Penelitian

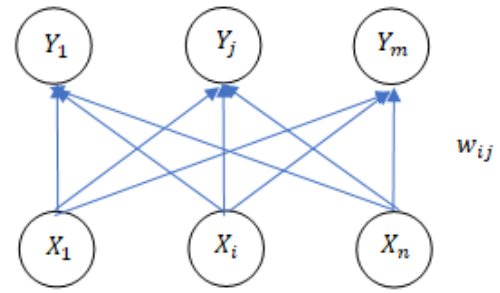
Metode yang digunakan pada penelitian adalah aplikasi Neural Network yang digunakan pada masalah klasifikasi yaitu Learning Vector Quantization (LVQ) dan Backpropagation (BP).

2.1 Learning Vector Quantization

Learning Vector Quantization (LVQ) ditemukan oleh Kohonen pada tahun 1990. Berdasarkan modifikasi pada bobot, terdapat dua jenis pelatihan (training) yaitu supervised dan unsupervised training. Pada supervised training, terdapat pasangan input dan target yang digunakan untuk melatih jaringan sampai matriks bobot optimal diperoleh. Untuk setiap training, inputs diberikan pada jaringan. Jaringan akan memproses output. Error antara output dan target dipropagasikan untuk mengupdate matriks bobot [1].

Pada unsupervised training, terdapat vektor bobot. Vektor bobot untuk suatu unit output pada jaringan berperan sebagai vektor untuk pola input. Selama training, suatu jaringan menentukan unit output yang terbaik untuk vektor input. Vektor bobot untuk pemenang disesuaikan menurut algoritma pada jaringan [1].

Model LVQ adalah model supervised training yang dapat dilihat pada Gambar 1.



Gambar 1. Model Jaringan Learning Vector Quantization (LVQ)

Pada LVQ, terdapat beberapa parameter, yaitu :

- N : jumlah data
- M : jumlah klasifikasi
- x_i : data ke- i
- c_i : klasifikasi pada data ke- i
- w_j : bobot dari klasifikasi ke- j

Algoritma LVQ adalah sebagai berikut :

Inisialisasi matriks bobot W sebagai vektor referensi.

$e = 1$

while($e \leq \max_epoch$ & & $dist \geq \min_dist$)

for($i = 1 : N$)

1. Hitung jarak euclidean $dist(j)$

$$dist(j) = \|w_j - x_i\| \quad j = 1, 2, \dots, M \quad (1)$$

2. Cari index J sehingga $dist(j)$ minimum.

3. Update $w_{j(\min)}$ dengan :

If $C_i = j(\min)$

$$w_{j(\min)}^{new} = w_{j(\min)}^{old} + \alpha(x_i - w_{j(\min)}^{old}) \quad (2)$$

Else

$$w_{j(\min)}^{new} = w_{j(\min)}^{old} - \alpha(x_i - w_{j(\min)}^{old}) \quad (3)$$

End

end

Hitung jarak euclidean pada persamaan (4)

$$dist = \sum_{d=1}^{datasize} \|x_d - w^*\| \quad (4)$$

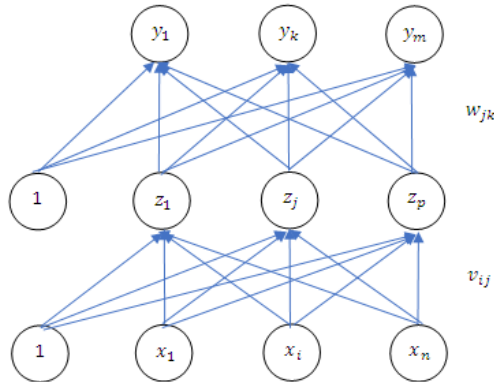
Turunkan learning rate $\alpha = \delta\alpha$ dengan δ adalah bilangan antara 0-1.

$e = e + 1$

end

2.2 Backpropagation

Backpropagation (BP) adalah jenis Neural Network yang digunakan pada proses estimasi. Model BP dapat dilihat pada Gambar 2.



Gambar 2. Model Jaringan Backpropagation (BP)

BP terdiri dari beberapa input x_1, x_2, \dots, x_n , beberapa hidden layer z_1, z_2, \dots, z_p , dan beberapa output y_1, y_2, \dots, y_m . Pada input dan hidden layer, terdapat bias yang memiliki nilai 1. Bobot v_{ij} menghubungkan input x_i ke hidden layer z_j . Bobot w_{jk} menghubungkan hidden layer z_j ke output y_k . Pada BP, terdapat tiga fase perhitungan yaitu forward propagation, backward propagation, dan update matriks bobot [1].

Pada Backpropagation (BP), terdapat fungsi aktivasi yang berfungsi menentukan output dari neuron. Argument dari fungsi aktivasi adalah kombinasi linear dari input dan bobot.

$$\begin{aligned} net &= \sum_i x_i w_i \\ f(net) &= f\left(\sum_i x_i w_i\right) \end{aligned} \quad (5)$$

Sifat dari fungsi aktivasi adalah kontinu, differentiable, dan bukan fungsi turun [1]. Pada penelitian ini, fungsi aktivasi yang dipakai adalah binary sigmoid dengan range (0-1) seperti pada persamaan (6).

$$\begin{aligned} f(net) &= \frac{1}{1 + e^{-net}} \\ f'(net) &= f(net)(1 - f(net)) \end{aligned} \quad (6)$$

Algoritma BP adalah sebagai berikut :

Inisialisasi matriks bobot V and W dengan bilangan random antara -0.5 sampai 0.5.

$e = 1$

while($e \leq \max_epoch$ & & $MSE \geq \min_MSE$)

for($d = 1 : \text{datasize}$)

1. Masing-masing input menerima sinyal dan melanjutkan ke setiap hidden layer melalui

forward propagation pada persamaan (7) sampai persamaan (10) dan backward propagation pada persamaan (11) sampai persamaan (15).

Forward Propagation

2. Hitung semua output $z_j, j = 1, 2, \dots, p$ pada hidden layer

$$z_net_j = v_{oj} + \sum_{i=1}^n x_i v_{ij} \quad (7)$$

$$z_j = f(z_net_j) = \frac{1}{1 + e^{-z_net_j}} \quad (8)$$

3. Hitung semua output $y_k, k = 1, 2, \dots, m$ pada output layer

$$y_net_k = w_{ok} + \sum_{j=1}^p z_j w_{jk} \quad (9)$$

$$y_k = f(y_net_k) = \frac{1}{1 + e^{-y_net_k}} \quad (10)$$

Backward Propagation

4. Hitung factor δ output berdasarkan error pada setiap output layer $y_k, k = 1, 2, \dots, m$

$$\delta_k = (t_k - y_k) f'(y_net_k) \quad (11)$$

5. Hitung perubahan bobot

$$\begin{aligned} \Delta w_{jk} &= \alpha \delta_k z_j, \\ k &= 1, 2, \dots, m \quad j = 0, 1, 2, \dots, p \end{aligned} \quad (12)$$

6. Hitung factor δ hidden layer berdasarkan error pada setiap hidden layer $z_j, j = 1, 2, \dots, p$

$$\delta_net_j = \sum_{k=1}^m \delta_k w_{jk} \quad (13)$$

$$\delta_j = \delta_net_j f'(z_net_j) \quad (14)$$

7. Hitung perubahan bobot

$$\begin{aligned} \Delta v_{ij} &= \alpha \delta_j x_i, \\ j &= 1, 2, \dots, p \quad i = 0, 1, 2, \dots, n \end{aligned} \quad (15)$$

Update Matriks Bobot

8. Update matriks bobot yang baru

$$w_{jk} = w_{jk} + \Delta w_{jk} \quad (16)$$

$$v_{ij} = v_{ij} + \Delta v_{ij} \quad (17)$$

end

De-normalisasi output untuk membandingkan dengan target asli

Hitung Root of Mean Square Error (RMSE)

$$RMSE = \sqrt{\frac{1}{datasize} \sum_{d=1}^{datasize} \frac{1}{m} \sum_{k=1}^m (T_{dk} - Y_{dk})^2} \quad (18)$$

with T_{dk} is target value and Y_{dk} is outputs.

$e = e + 1$

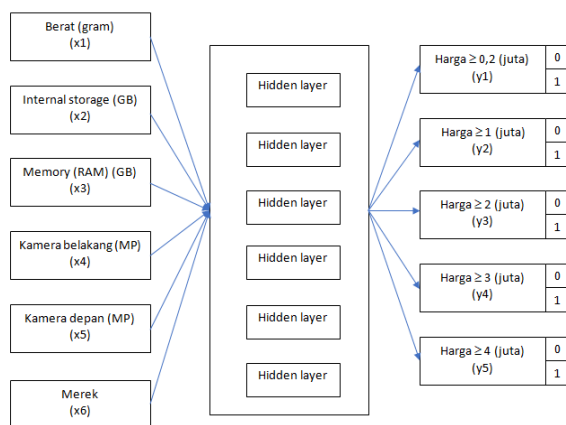
end

2.3 Konstruksi pada Data

Pada klasifikasi harga smartphone di pasaran, terdapat dataset dari smartphone, dimana atribut yang digunakan sebagai input adalah : berat (gram), penyimpanan internal (GB), memory (RAM) (GB), kamera belakang (MP), kamera depan (MP) dan merek seperti pada Gambar 3. Atribut tersebut akan digunakan sebagai input pada algoritma BP untuk menghasilkan output, yaitu kelas harga smartphone. Sedangkan pada LVQ, akan diambil sampel tiap kelas sebagai vektor reference. Kelas harga smartphone yang digunakan adalah :

1. Kelas 1 (harga smartphone di atas Rp. 200.000)
2. Kelas 2 (harga smartphone di atas Rp. 1.000.000)
3. Kelas 3 (harga smartphone di atas Rp. 2.000.000)
4. Kelas 4 (harga smartphone di atas Rp. 3.000.000)
5. Kelas 5 (harga smartphone di atas Rp. 4.000.000)

Pada klasifikasi multi output, pada jaringan BP terdapat layer output lebih dari satu seperti pada Gambar 3, yaitu harga di atas Rp. 200.000, harga di atas Rp. 1.000.000, harga di atas Rp. 2.000.000, harga di atas Rp. 3.000.000, dan harga di atas Rp. 5.000.000.



Gambar 3. Konstruksi Data pada Jaringan Backpropagation (BP)

Satu obyek bisa memiliki lebih dari satu output. Contoh smarphone harga Rp. 2.500.000, masuk ke dalam Kelas 3, yaitu di atas Rp.

2.000.000. Namun juga pasti masuk dalam Kelas 1 (di atas Rp. 200.000) dan Kelas 2 (di atas Rp. 1.000.000), tapi tidak untuk Kelas 4 (di atas Rp. 3.000.000) dan Kelas 5 (di atas Rp. 4.000.000), sehingga output dari smarphone tersebut adalah [11100].

Pada klasifikasi, performace yang digunakan adalah accuracy, error rate, recall, specificity, dan precision dengan rumus sebagai berikut [2] :

$$accuracy = \frac{TP + TN}{P + N} \times 100\% \quad (19)$$

$$error\ rate = \frac{FP + FN}{P + N} \times 100\% \quad (20)$$

$$recall = \frac{TP}{P} \times 100\% \quad (21)$$

$$specificity = \frac{TN}{N} \times 100\% \quad (22)$$

$$precision = \frac{TP}{TP + FP} \times 100\% \quad (23)$$

Dengan penjelasan :

TP : Jumlah tuple positif yang diprediksi benar sebagai positif oleh classifier

TN : Jumlah tuple negatif yang diprediksi benar sebagai negatif oleh classifier

FP : Jumlah tuple negatif yang diprediksi salah sebagai positif oleh classifier

FN : Jumlah tuple positif yang diprediksi salah sebagai negatif oleh classifier

P : Jumlah tuple positif pada data target

N : Jumlah tuple negatif pada data target

3. Hasil Dan Pembahasan

Data smartphone yang digunakan diperoleh dari daftar smartphone yang dijual di pasaran pada Februari - Maret 2020 dengan 80 jenis smartphone. Kemudian klasifikasi ke dalam Kelas 1 (harga smartphone di atas Rp. 200.000), Kelas 2 (harga smartphone di atas Rp. 1.000.000), Kelas 3 (harga smartphone di atas Rp. 2.000.000), Kelas 4 (harga smartphone di atas Rp. 3.000.000), dan Kelas 5 (harga smartphone di atas Rp. 4.000.000)

Sebelum menerapkan estimasi dan proses klasifikasi, bagi data ke dalam data training dengan proporsi 80% dari semua data dan data testing dengan proporsi 20% sisanya. Kemudian normalisir data sehingga berada di antara 0-1 [2].

3.1 Klasifikasi pada Learning Vector Quantization

Parameter yang digunakan pada simulasi LVQ adalah :

Learning rate α : 0.6

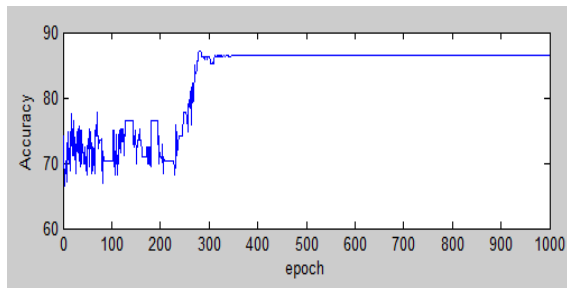
Learning rate δ : 0.9

Maximum epoch : 1000

Simulasi LVQ pada klasifikasi kelas harga smartphone dapat dilihat pada Gambar 4. Pertama-tama, inialisasi matriks bobot dari lima kelas :

1. Kelas 1 (harga smartphone di atas Rp. 200.000)
2. Kelas 2 (harga smartphone di atas Rp. 1.000.000)
3. Kelas 3 (harga smartphone di atas Rp. 2.000.000)
4. Kelas 4 (harga smartphone di atas Rp. 3.000.000)
5. Kelas 5 (harga smartphone di atas Rp. 4.000.000)

Gunakan pada data training menggunakan LVQ sampai maksimum epoch dan proses konvergensi dapat dilihat pada Gambar 4. Terlihat pada awal epoch, accuracy pada persamaan (19) yang dihasilkan cukup kecil. Pada proses optimisasi, accuracy meningkat dan konvergen dengan accuracy 86.5625 % pada data training. Kemudian matriks bobot optimal pada data training digunakan pada data testing dan menghasilkan accuracy 71.25 % pada data testing.



Gambar 4. Proses Optimisasi Accuracy Algoritma LVQ pada Data Training

Sehingga performance dari algoritma LVQ pada data training dan data testing seperti pada Tabel 1.

Tabel 1. Performance Klasifikasi pada Algoritma LVQ

Training Data (64 data)	
Accuration	86,5625 %
Error rate	13,4375 %
Recall	93,4884 %
Specificity	72,3810 %
Precision	87,3913 %
Testing Data (16 data)	
Accuration	71,25 %
Error rate	28,75 %
Recall	68,5185 %
Specificity	76,9231 %
Precision	86,0465 %

3.2 Klasifikasi pada Backpropagation

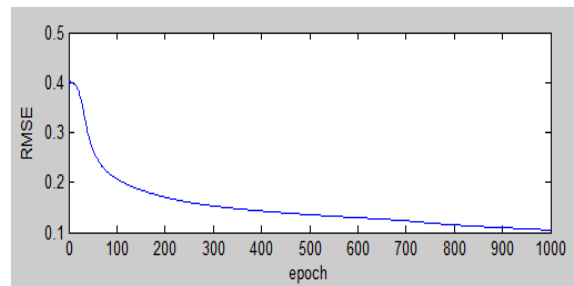
Parameter yang digunakan pada simulasi BP adalah :

Learning rate α : 0.2

Jumlah hidden layer : 2
Maximum epoch : 1000

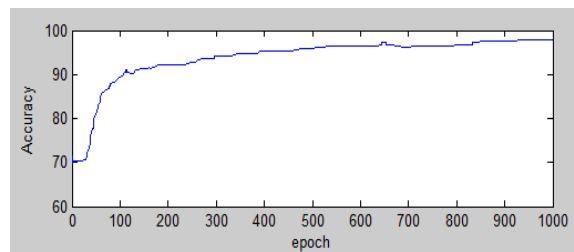
Simulasi BP pada klasifikasi kelas harga smartphome dapat dilihat pada Gambar 5 pada optimisasi nilai RMSE dan Gambar 6 pada optimisasi nilai accuracy.

Pada optimisasi nilai RMSE, inialisasi matriks bobot dan gunakan pada training data menggunakan BP sampai maksimum epoch dan proses konvergensi dapat dilihat pada Gambar 5. Terlihat pada awal epoch, RMSE pada persamaan (18) yang dihasilkan cukup besar. Pada proses optimisasi, RMSE menurun dan konvergen dengan nilai RMSE adalah 0,1048 pada data training. Kemudian matriks bobot optimal pada data training digunakan pada data testing dan menghasilkan RMSE 0,2293 pada data testing.



Gambar 5. Proses Optimisasi RMSE Algoritma BP pada Data Training

Pada optimisasi nilai accuracy, inialisasi matriks bobot dan gunakan pada training data menggunakan BP sampai maksimum epoch dan proses konvergensi dapat dilihat pada Gambar 6. Terlihat pada awal epoch, accuracy pada persamaan (19) yang dihasilkan cukup kecil. Pada proses optimisasi, accuracy meningkat dan konvergen dengan nilai accuracy adalah 97,8125 % pada data training. Kemudian matriks bobot optimal pada data training digunakan pada data testing dan menghasilkan accuracy 90 % pada data testing.



Gambar 6. Proses Optimisasi Accuracy Algoritma BP pada Data Training

Sehingga performance dari algoritma BP pada data training dan data testing seperti pada Tabel 2.

Tabel 2. Performance Klasifikasi pada Algoritma BP

Training Data (64 data)	
Accuration	97,8125 %
Error rate	2,1875 %
Recall	96,7442 %
Specificity	100 %
Precision	100 %
Testing Data (16 data)	
Accuration	90 %
Error rate	10 %
Recall	88,8889 %
Specificity	92,3077 %
Precision	96 %

4. Kesimpulan

Terdapat dua metode yang digunakan pada klasifikasi kelas harga smartphone di pasaran yaitu Learning Vector Quantization (LVQ) dan Backpropagation (BP). Pada proses klasifikasi menggunakan LVQ dan BP, terdapat perbedaan pada kedua metode tersebut. LVQ mengklasifikasi harga smartphone menggunakan jarak euclidean antara bobot (weight) dengan data di setiap iterasi. BP mengklasifikasi harga smartphone menggunakan penurunan gradien (gradient descent) dari target dan output di setiap iterasi. Berdasarkan hasil simulasi, BP memberikan accuracy yang lebih baik pada data training maupun data testing daripada LVQ. Pengembangan dari penelitian ini adalah teknik klasifikasi dengan big data menggunakan machine learning process.

Daftar Pustaka

- [1] L. Fausett, *Fundamental of Neural Networks*, Prentice Hall, New York, 1994
- [2] J. Han, M. Kamber, J. Pei, *Data Mining Concepts and Techniques*, Elsevier, New York, 2012
- [3] D. Rahmalia, T. Herlambang, "Prediksi Cuaca Menggunakan Algoritma Particle Swarm Optimization-Neural Network (PSO-NN)" *In Prosiding Seminar Nasional Matematika dan Aplikasinya*, pp. 41-48, 2017
- [4] D. Rahmalia, T. Herlambang, "Application Kohonen Network and Fuzzy C Means for Clustering Airports Based on Frequency of Flight," *Kinetik : Game Technology, Information System, Computer Network, Computing* ; vol. 3, no. 3, pp. 229-236. 2018.
- [5] A. Muhith, T. Herlambang, Irhamah, D. Rahmalia, "Estimation of Thrombocyte Concentrate (TC) and Whole Blood (WB) Using Unscented Kalman Filter," *International Journal of Advanced Science and Technology* ; vol. 9, no. 8. 2020.
- [6] D. Rahmalia, N. Aini, "Pengaruh Korelasi Data pada Peramalan Suhu Udara Menggunakan Backpropagation Neural Network," *Zeta Math Journal* ; vol. 4, no. 1, pp.1-6. 2018.
- [7] D. Rahmalia, M.S. Pradana, "Backpropagation Neural Network pada Data yang Tak Stationer," *Jurnal Riset dan Aplikasi Matematika (JRAM)* ; vol. 3, no. 1, pp. 32-42. 2019.
- [8] A. Rohmatullah, D. Rahmalia, M.S. Pradana, "Klasterisasi Data Pertanian di Kabupaten Lamongan Menggunakan Algoritma K-Means dan Fuzzy C Means," *Jurnal Ilmiah Teknosains* ; vol. 5, no. 2, pp. 86-93. 2020.
- [9] D. Rahmalia, A. Rohmatullah, "Pengaruh Korelasi Data pada Peramalan Kelembaban Udara Menggunakan Adaptive Neuro Fuzzy Inference System (ANFIS)," *Applied Technology and Computing Science Journal* ; vol. 2, no. 1, pp. 10-24. 2019.
- [10] D.F. Karya, P. Katias, T. Herlambang, D. Rahmalia, "Development of Unscented Kalman Filter Algorithm for Stock Price Estimation" *In Journal of Physics : Conference Series*, 2019
- [11] D. Rahmalia et al, "Comparison Between Neural Network (NN) and Adaptive Neuro Fuzzy Inference System (ANFIS) on Sunlight Intensity Prediction Based on Air Temperature and Humidity" *In Journal of Physics : Conference Series*, 2019