

# Integrasi Metode K-Means Clustering Dan Backpropagation Pada Pemodelan Peramalan Penjualan

Mochamad Zamzamik<sup>1</sup>, Daniel Swanjaya<sup>2</sup>

Teknik Informatika, Fakultas Teknik, Universitas Nusantara PGRI Kediri

E-mail: [1zam.zamik777@gmail.com](mailto:1zam.zamik777@gmail.com), [2daniel@unpkediri.ac.id](mailto:2daniel@unpkediri.ac.id)

**Abstrak** – Model peramalan penjualan dibutuhkan untuk membangun aplikasi peramalan penjualan. Banyak penelitian telah dilakukan, salah satunya model peramalan menggunakan metode Jaringan Syaraf Tiruan Backpropagation, tetapi kualitas model peramalannya belum memuaskan. Penelitian ini bertujuan untuk memperbaiki kualitas peramalan dengan cara mengintegrasikan metode K-Means Clustering dan Backpropagation. Dataset yang digunakan adalah data penjualan Bolu Pisang di Salsabila Cake. Proses diawali dengan transformasi data atau normalisasi, kemudian pengelompokan menggunakan K-Means Clustering, berikutnya pada setiap kelompok dilakukan pelatihan dan uji coba menggunakan Backpropagation, output dari Backpropagation dinormalisasi untuk mendapatkan nilai yang sesungguhnya. Kualitas peramalan ditentukan menggunakan Mean Absolute Deviation (MAD), dengan menghitung rata-rata selisih output dengan nilai nyata. Pada penelitian ini didapati nilai Silhouette pengelompokan yang terbaik didapat dari Vektor Fitur yang panjangnya 16 dan banyak kelompoknya 12, sebesar 0.0349. Kemudian nilai rata-rata MAD dari semua kelompok adalah 5.5 dan 4.5 untuk proses pelatihan dan uji coba, hasil ini lebih baik daripada penelitian sebelumnya yang hanya menggunakan Backpropagation saja untuk peramalan penjualan.

**Kata Kunci** — Backpropagation, K-Means, MAD

## 1. PENDAHULUAN

Jaringan Syaraf Tiruan (JST) atau yang sering dikenal dengan *Neural Network* (NN) adalah sistem pemrosesan informasi yang memiliki karakteristik mirip dengan jaringan syaraf biologis. NN merupakan salah satu dari sistem informasi yang didesain dengan menirukan kerja otak manusia dalam menyelesaikan suatu masalah dengan melakukan proses belajar melalui perubahan bobot sinapsisnya. NN telah dikembangkan sebagai generalisasi model matematika dari kognisi manusia atau syaraf biologi. NN ditandai dengan pola hubungan antara neuron (arsitektur), algoritma untuk menentukan bobot penghubung (pelatihan, atau belajar, algoritma), dan fungsi aktivasi [1]. NN bermanfaat untuk pengenalan pola, signal processing, pengklasifikasian dan peramalan [2].

Peramalan penjualan adalah kegiatan untuk mengestimasi besarnya penjualan barang atau jasa oleh produsen, distributor pada periode waktu dan wilayah pemasaran tertentu. Peramalan penjualan merupakan bagian fungsi manajemen sebagai salah satu kontributor keberhasilan sebuah perusahaan. Ketika penjualan diprediksi dengan akurat maka pemenuhan permintaan konsumen dapat diusahakan tepat waktu, kerjasama perusahaan dengan relasi tetap terjaga dengan baik, kepuasan konsumen terpenuhi, perusahaan dapat mengatasi hilangnya penjualan atau kehabisan stok, mencegah pelanggan lari ke kompetitor. Di sisi lain perusahaan dapat menentukan keputusan kebijakan rencana produksi,

persediaan barang, investasi aktiva dan *cash flow*. Dengan kata lain, tidak ada perusahaan yang dapat menghindar dari kegiatan memperkirakan atau meramalkan penjualan untuk keperluan perencanaan aktivitas-aktivitas yang harus dilakukan [3].

Berdasarkan kemampuan belajar yang dimilikinya, maka NN dapat dilatih untuk mempelajari dan menganalisa pola data masa lalu dan berusaha mencari suatu formula atau fungsi yang akan menghubungkan pola data masa lalu dengan output yang diinginkan pada saat ini [4]. Salah satu kelebihan dari model NN dalam algoritma peramalan yaitu dapat digunakan untuk meramalkan data time series non linier [5].

Suhada (2009) dalam penelitiannya menyimpulkan bahwa peramalan dengan menggunakan pendekatan NN memberikan nilai MSE yang sangat kecil mendekati nol. Hal ini berarti bahwa NN memiliki akurasi yang sangat tinggi dalam melakukan peramalan terhadap suatu model sistem [6]. Menurut Walid, dkk (2015), peramalan menggunakan NN menghasilkan hasil yang jauh lebih baik dibandingkan dengan model ARIMA (*Autoregressive Integrated Moving Average*) [7]. Kemudian menurut Kohzadi, dkk (1996) dalam penelitian mereka mengungkapkan bahwa peramalan dengan NN jauh lebih akurat dibandingkan dengan model tradisional ARIMA. Alasan mengapa model NN lebih baik daripada model ARIMA karena data mengandung perilaku non linier yang tidak dapat sepenuhnya ditangkap oleh model ARIMA linier [8].

Namun, jika dibandingkan dengan algoritma peramalan tradisional seperti model ARIMA atau model regresi, ada banyak faktor pemodelan yang perlu dipertimbangkan dalam NN [9].

Penelitian yang menggunakan *Backpropagation* untuk peramalan diantaranya adalah, penelitian yang dilakukan oleh Adi Sukarno Rachman (2017) dengan judul “Peramalan Produksi Gula Menggunakan Metode Jaringan Syaraf Tiruan *Backpropagation* Pada PG Candi Baru Sidoarjo”. Keberhasilan dari penelitian tersebut didapatkan nilai *MAPE* (*Mean Absolute Percentage Error*) terendah sebesar 17,85% dengan jumlah iterasi 800. Dan pada pengujian *learning rate* didapatkan nilai *MAPE* terendah sebesar 17,38% dengan nilai *learning rate* 0,4. Jika dengan nilai iterasi maksimum 800 dan nilai *learning rate* 0,4 maka akan menghasilkan nilai *MAPE* sebesar 16,98% [10].

Penelitian lainnya yang dilakukan oleh M. Ihsan Diputra (2017) dengan judul “Prediksi Tingkat Keuntungan Usaha Peternakan Itik Alabio Petelur Menggunakan Jaringan Syaraf Tiruan *Backpropagation* (Kasus Di Kabupaten Hulu Sungai Utara Kalimantan Selatan)”. Dari keberhasilan pengujian *learning rate* didapatkan nilai parameter yang optimal. Nilai parameter tersebut yaitu *learning rate* sebesar 0,8, jumlah neuron pada hidden layer sebesar 17, batas *MAPE* latih sebesar 10%, dan jumlah data latih sebanyak 90% dari total data latih. Nilai *MAPE* terbaik yang didapatkan sebesar 25,7852% [11].

Kemudian penelitian yang dilakukan oleh Muhammad Imam Ghozali (2016) dengan judul “Rantai Pasok Beras Pada Bulog Berbasis Neural Network”. Dari penelitian tersebut didapatkan data prediksi beras untuk tahun berikutnya dengan menggunakan Neural Network Algoritma *Backpropagation* dengan *learning rate* 0,8 dan momentum 0,2 [12]

Berdasarkan latar belakang tersebut didapati kualitas peramalan pada penelitian sebelumnya masih belum optimal, maka penulis membuat penelitian yang berjudul “Integrasi Metode *K-Means Clustering* dan *Backpropagation* Pada Pemodelan Peramalan Penjualan”. Penelitian ini bertujuan untuk memperbaiki kualitas Peramalan dengan menambahkan proses clustering sebelum dilakukan proses peramalan, yaitu dengan menggunakan metode *K-Means*. Metode *K-Means* dipilih karena sederhana untuk diimplementasikan atau digunakan, selain itu metode *K-Means* dapat dijalankan dengan relatif cepat, mudah beradaptasi, dan juga sudah umum penggunaannya. Selanjutnya Proses peramalan dengan menggunakan metode

*Backpropagation* untuk tiap kelompok yang sudah terbentuk.

## 2. METODE PENELITIAN

### 2.1 Metode K-Means

*K-Means* merupakan metode non hierarki yang pada awalnya mengambil sebagian banyaknya komponen populasi untuk dijadikan pusat cluster awal. Pusat cluster dipilih secara acak (*random*) dari sekumpulan data, berikutnya *K-Means* akan menguji dari masing-masing komponen yang ada didalam data dan akan menandai komponen tersebut kesalah satu pusat cluster yang telah di definisikan, tergantung dari jarak minimum antar komponen dengan tiap-tiap cluster. Posisi pusat cluster dihitung akan kembali sampai semua komponen data digolongkan ke dalam tiap-tiap pusat cluster dan yang terakhir akan terbentuk posisi pusat cluster yang baru [13].

Algoritma *K-Means* mampu mengcluster data besar dengan sangat cepat. Dalam Algoritma *K-Means*, setiap data harus masuk ke dalam cluster tertentu dan bisa dimungkinkan bagi setiap data yang termasuk ke dalam cluster tertentu pada suatu tahap proses. Pada tahap proses berikutnya berpindah ke dalam cluster lainnya. Berikut ini merupakan aturan yang ada di dalam Algoritma *K-Means* [14]

1. Yang pertama yaitu memasukan jumlah cluster
2. Algoritma *K-Means* hanya memiliki atribut yang bertipe numerik.

Berikut ini merupakan langkah-langkah untuk pengelompokan data :

1. Pilih jumlah cluster
2. Inisialisasi awal dan pusat cluster dilakukan secara acak
3. Setiap data di tempatkan ke pusat cluster terdekat berdasarkan jarak antar objek. Jarak dihitung dengan menentukan kemiripan dan ketidak miripan data dengan menggunakan Metode Jarak Euclidean (*Euclidean Distance*) dengan rumus persamaan 1.

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \dots\dots\dots(1)$$

Dimana :

$d(x, y)$  = ukuran ketidak miripan

$x_i$  =  $(x_1, x_2, x_3, \dots, x_i)$  variabel data

$y_i$  =  $(y_1, y_2, y_3, \dots, y_i)$  variabel yang ada pada titik pusat

### 2.2 Backpropagation

*Backpropagation* merupakan *Jaringan Syaraf Tiruan* (JST) dengan layer tunggal yang memiliki kelemahan pada keterbatasan dalam pengenalan pola.

Dengan menambahkan satu atau beberapa layer tersembunyi diantara layer masukan dan layer keluaran, merupakan cara yang digunakan untuk mengatasi kelemahan tersebut. *Backpropagation* dapat melatih jaringan untuk mendapatkan keseimbangan antara kemampuan jaringan untuk mengenali suatu pola yang akan digunakan selama dalam pelatihan, serta mempunyai kemampuan dalam memberi respon yang benar terhadap suatu pola masukan yang serupa dengan pola yang digunakan.

Berikut ini merupakan penjelasan dari tahap-tahap untuk melakukan pelatihan *Backpropagation* [15] :

1. Algoritma *Backpropagation*
  1. menginisialisasi bobot awal (pilih nilai secara acak yang cukup kecil)
  2. Selama kondisi masih berhenti, maka bernilai salah
2. Tahap perambatan-maju (*Forward Propagation*)
  1. Pada setiap unit input ( $x_i, i = 1, 2, 3, \dots, n$ ) akan menerima sinyal  $x_i$  lalu akan meneruskan sinyal tersebut ke semua unit yang terdapat pada lapisan tersembunyi
  2. Pada setiap unit tersembunyi ( $z_i, i = 1, 2, 3, \dots, p$ ) akan menjumlahkan bobot sinyal input dengan menggunakan persamaan 2.
 
$$z\_in_j = v_{0j} + \sum_{i=1}^n x_i v_{ij} \dots\dots\dots (2)$$
  3. Lalu menerapkan fungsi aktivasi yang berguna untuk menghitung sinyal outputnya menggunakan persamaan 3.
 
$$z_i = f(z\_in_j) \dots\dots\dots (3)$$
  4. Fungsi aktivasi yang biasa digunakan yaitu fungsi *sigmoid*, kemudian mengirimkan sinyal tersebut ke semua unit *output*
  5. Pada setiap unit input ( $y_{ki}, k = 1, 2, 3, \dots, m$ ) akan menjumlahkan bobot sinyal *input* dengan menggunakan persamaan 4.
 
$$y\_in_k = w_{0k} + \sum_{i=1}^p z_{iw} w_{jk} \dots\dots\dots (4)$$
  6. Selanjutnya menerapkan fungsi aktivasi yang berguna untuk menghitung sinyal *output* dengan menggunakan persamaan 5.
 
$$y_k = f(y\_in_k) \dots\dots\dots (5)$$
3. Tahap perambatan-balik (*Backpropagation*)
  1. Pada setiap unit *output* ( $y_k, k = 1, 2, 3, \dots, m$ ) akan menerima pola target yang sesuai dengan pola input pelatihan, kemudian menghitung *error* dengan menggunakan persamaan 6.
 
$$\delta_k = (t_k - y_k) f'(y\_in_k) \dots\dots\dots (6)$$
  2.  $F'$  adalah turunan dari aktivasi, kemudian menghitung koreksi bobot dengan menggunakan persamaan 7.
 
$$\Delta w_{jk} = a \delta_j z_j \dots\dots\dots (7)$$

3. Selanjutnya menghitung koreksi bias dengan menggunakan persamaan 8.
 
$$\Delta w_{0j} = a \delta_k \dots\dots\dots (8)$$
4. Selanjutnya akan mengirimkan ke unit-unit yang ada dilapisan tersembunyi ( $y_{i,i} = 1, 2, 3, \dots, p$ ) setelah itu menjumlahkan data input-inputnya (dari unit-unit yang berada pada lapisan yang di kanannya) dengan menggunakan persamaan 9.
 
$$\delta\_in_j = \sum_{k=1}^m \delta_k w_{jk} \dots\dots\dots (9)$$
5. Selanjutnya menghitung informasi *error*, dengan cara dikalikan dari fungsi aktivasinya dengan menggunakan persamaan 10.
 
$$\delta_j = \delta\_in_j f'(z - in_j) \dots\dots\dots (10)$$
6. Tahap selanjutnya menghitung koreksi bobot dengan menggunakan persamaan 11.
 
$$\Delta v_{jk} = a \delta_j x_i \dots\dots\dots (11)$$
7. Lalu menghitung koreksi bias dengan menggunakan persamaan 12.
 
$$\Delta v_{jk} = a \delta_j \dots\dots\dots (12)$$

### 2.3 MAD

*Mean Absolute Deviation* atau lebih sering disebut dengan MAD adalah perhitungan yang digunakan untuk menghitung rata-rata pada kesalahan yang mutlak. MAD digunakan untuk mengevaluasi metode peramalan menggunakan jumlah dari kesalahan-kesalahan. MAD mengukur ketepatan peramalan dengan merata-rata kesalahan.

Untuk mengukur tingkat keakuratan peramalan menggunakan metode *Mean Absolute Deviation* (MAD) dihitung menggunakan persamaan 13.

$$MAD = \frac{\sum |y_i - \hat{y}_i|}{n} \dots\dots\dots (13)$$

Keterangan	
MAD	: Mean Absolute Deviation
$y$	: Nilai Hasil Aktual
$\hat{y}$	: Nilai Hasil Prediksi
$n$	: Jumlah Data

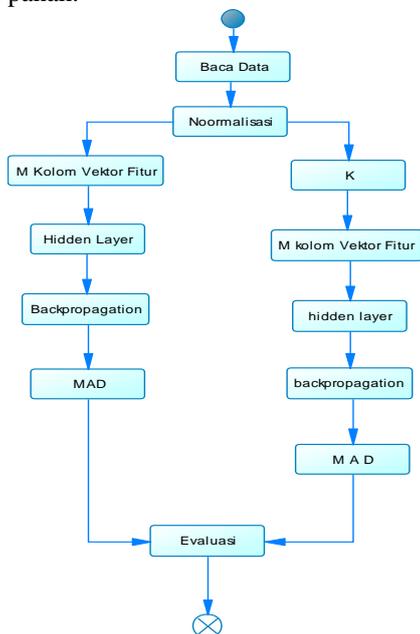
### 2.4 Perancangan Sistem

Pada bagian ini berisi sebuah perancangan yang ada pada suatu sistem yang akan dibuat untuk menghasilkan suatu sistem yang akan di implementasikan atau suatu sistem yang akan dijalankan. Berikut ini merupakan penjelasan dari perancangan sistem yang akan dibuat :

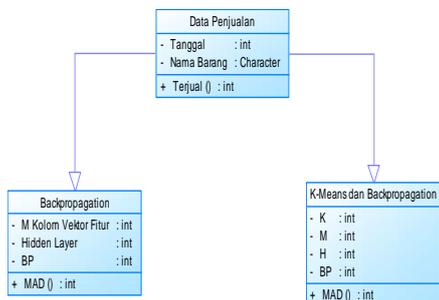
#### 2.4.1 Activity Diagram

Activity diagram merupakan rancangan aliran kerja atau aktivitas alur kerja pada sebuah sistem yang akan dijalankan. Activity diagram dapat berguna untuk mendefinisikan

suatu sistem atau mengelompokkan alur dari suatu sistem. Activity diagram mempunyai komponen yang berbentuk tertentu yang dihubungkan dengan menggunakan tanda panah.



Gambar 1 Activity Diagram

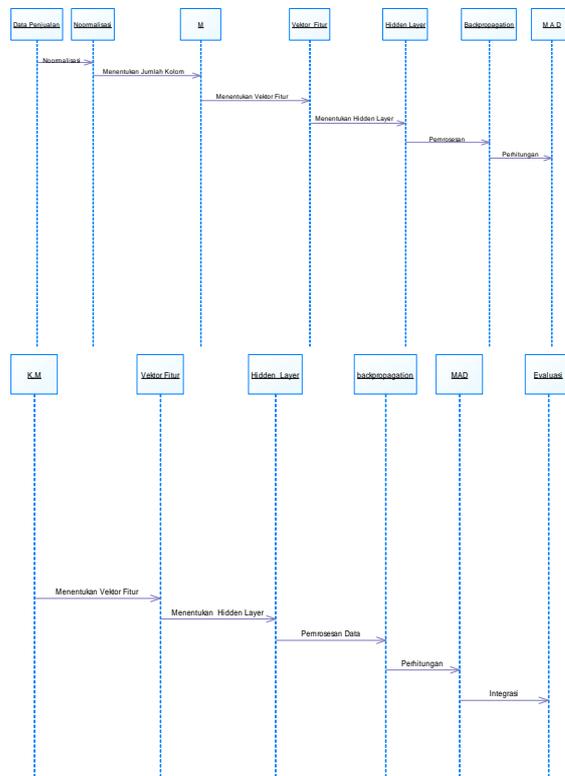


Gambar 2 Class Diagram

Pada gambar 1 dapat dijelaskan akan terdapat dua skenario percobaan untuk perbandingan yang akan di lakukan. Pertama, data yang sudah terbaca kemudian akan di noormalisasi yang kemudian baru dapat diproses dengan skenario atau percobaan yang pertama. Proses skenario atau tahapan yang pertama yaitu menentukan M, kolom Vektor Fitur, Hidden Layer. Setelah sudah di tentukan, kemudian di proses menggunakan perhitungan algoritma Backpropagation, setelah itu menghitung kesalahan (*error*) menggunakan MAD dengan persamaan 13.

Selanjutnya data di proses dengan skenario atau percobaan perbandingan yang ke dua. Tahapan percobaan yang kedua, pertama menentukan K Cluster, M kolom Vektor Fitur, Dan Hidden layer. Setelah itu di proses menggunakan perhitungan algoritma

Backpropagation, setelah itu menghitung kesalahan (*error*) menggunakan MAD dengan persamaan 13. Yang kemudian hasil MAD dari kedua percobaan di evaluasi.



Gambar 3 Sequence Diagram

#### 2.4.2 Class Diagram

Class Diagram merupakan suatu alur sistem yang berjenis diagram yang bisa memetakan struktur-struktur yang terdapat pada suatu sistem tertentu dengan cara memodelkan kelas, atribut, operasi, dan juga hubungan antar objek satu dengan objek yang lainnya.

pada gambar 2 merupakan Class Diagram pada penerapan data mining pada pemodelan peramalan penjualan yang digunakan untuk merancang sebuah sistem yang akan digunakan. Pada gambar 2 dapat dijelaskan bahwa, akan di olah dengan dua percobaan untuk dibandingkan. Pada percobaan yang pertama menggunakan algoritma Backpropagation, dan pada percobaan yang kedua menggunakan algoritma K-Means dan Backpropagation. Kemudian kesalahan (*error*) dari kedua percobaan di ukur dengan MAD yang kamudia nilai MAD dari kedua percobaan yang sudah di dapatkan akan di evaluasi.

### 2.4.3 Sequence Diagram

Sequence Diagram merupakan suatu diagram yang menggambarkan interaksi objek dan mengindikasikan objek atau memberi petunjuk dan tanda antar objek-objek. Pada Sequence Diagram ini, objek-objek akan diatur dengan berdasarkan waktu. Semua objek-objek yang berkaitan dengan proses berjalannya operasi, akan diurutkan berdasarkan waktu terjadinya mulai dari kiri ke kanan secara berurutan.

Pada gambar 3 dapat dijelaskan, data penjualan di normalisasi data di olah dengan skenario pertama yaitu, menentukan M kolom Vektor Fitur, Hidden Layer lalu di proses menggunakan perhitungan algoritma Backpropagation. Lalu *error* di hitung menggunakan MAD. Pada proses skenario kedua, menentukan K, M kolom Vektor Fitur, dan hidden Layer terlebih dahulu. Lalu kemudian di proses menggunakan algoritma Backpropagation. Lalu *error* di hitung dengan menggunakan MAD. Setelah itu pada tahap yang terakhir yaitu, evaluasi pada nilai MAD yang sudah di dapatkan dari kedua skenario.

## 3. HASIL DAN PEMBAHASAN

Penelitian ini bertujuan untuk membuat model peramalan atau prediksi penjualan dengan mengintegrasikan metode K-Means dan Backpropagation. Untuk mencapai tujuan tersebut penulis menggunakan data penjualan Bolu Pisang dari Salsabila Cake tahun 2016 sampai dengan 2020. Adapun proses integrasi tersebut menggunakan langkah-langkah sebagai berikut :

### 3.1 Sumber Data

Dalam penelitian ini, dataset didapat dari Salsabila Cake yang merupakan rekap data penjualan harian Bolu Pisang selama 5 tahun dari tahun 2016 sampai dengan 2020. Pada dataset ini terdapat 1826 data penjualan. Tabel 1 merupakan contoh data rekap penjualan harian bulan Februari 2019.

### 3.2 Praproses

Sebelum data penjualan diolah, data tersebut perlu ditransformasi menggunakan metode min-max (persamaan 14) sehingga domain data penjualan menjadi 0,1 hingga 0,9, seperti tabel 2.

$$X_{new} = 0.1 + 0.8 \times \frac{x_{old} - \min}{\max - \min} \dots\dots\dots (14)$$

Setelah proses transformasi selesai, perlu dibuat Vektor Fitur, dimana data penjualan didistribusikan ke bentuk vektor dengan panjang M. Tabel 3 adalah Vektor Fitur dengan M = 5, dimana kolom ke-1 berisi

penjualan hari ke-1, kolom ke-2 berisi penjualan hari ke-2, hingga kolom ke-5 yang berisi penjualan hari ke-5.

Tabel 1. Contoh Data Penjualan Bolu Pisang

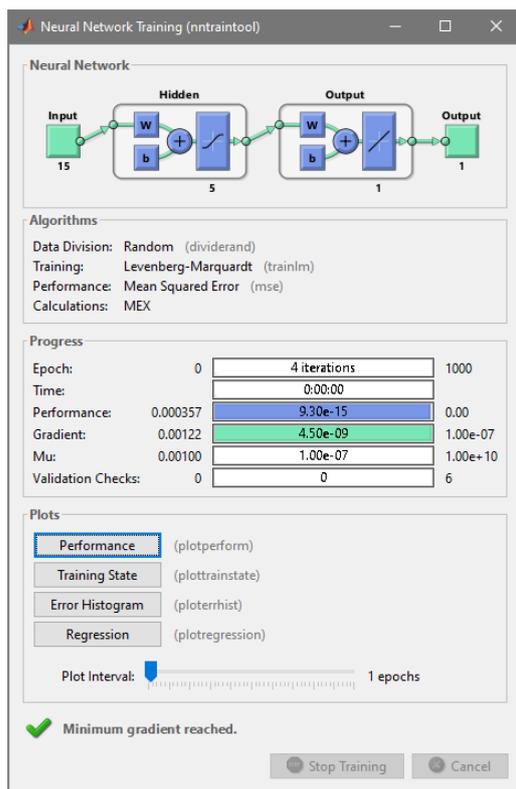
Tanggal	Banyak	Tanggal	Banyak
1	2	15	0
2	0	16	1
3	0	17	1
4	5	18	5
5	3	19	0
6	3	20	3
7	12	21	6
8	2	22	10
9	0	23	1
10	0	24	0
11	0	25	0
12	2	26	1
13	6	27	1
14	1	28	2

Tabel 2. Contoh Data yang sudah ditransformasi

Tanggal	Banyak	Tanggal	Banyak
1	0,1533	15	0,1000
2	0,1000	16	0,1267
3	0,1000	17	0,1267
4	0,2333	18	0,2333
5	0,1800	19	0,1000
6	0,1800	20	0,1800
7	0,4200	21	0,2600
8	0,1533	22	0,3667
9	0,1000	23	0,1267
10	0,1000	24	0,1000
11	0,1000	25	0,1000
12	0,1533	26	0,1267
13	0,2600	27	0,1267
14	0,1267	28	0,1533

Tabel 3. Contoh Vektor Fitur dengan M = 5

No	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>
1	0,1533	0,1000	0,1000	0,2333	0,1800
2	0,1000	0,1000	0,2333	0,1800	0,1800
3	0,1000	0,2333	0,1800	0,1800	0,4200
4	0,2333	0,1800	0,1800	0,4200	0,1533
5	0,1800	0,1800	0,4200	0,1533	0,1000
6	0,1800	0,4200	0,1533	0,1000	0,1000
7	0,4200	0,1533	0,1000	0,1000	0,1000
8	0,1533	0,1000	0,1000	0,1000	0,1533
9	0,1000	0,1000	0,1000	0,1533	0,2600
10	0,1000	0,1000	0,1533	0,2600	0,1267
11	0,1000	0,1533	0,2600	0,1267	0,1000
12	0,1533	0,2600	0,1267	0,1000	0,1267



Gambar 4 Tampilan Neural Network Training salah satu pengujian

Tabel 4. Hasil pengujian Silhouette Pengelompokan ( $\times 10^{-4}$ )

K	M				
	14	15	16	17	18
2	205	212	200	202	208
3	207	220	201	206	216
4	226	252	208	221	239
5	229	255	231	232	249
6	263	271	267	234	261
7	266	312	278	242	269
8	268	317	289	244	275
9	270	323	301	255	295
10	299	329	311	266	296
11	327	333	341	314	312
12	344	346	349	345	326

Tabel 5. Nilai MAD proses Pelatihan dan Ujicoba

Kelompok ke-	Pelatihan		Ujicoba	
	N	MAD	N	MAD
1	126	7	31	5
2	149	5	37	4
3	106	7	27	4
4	107	5	27	4
5	126	4	31	5
6	129	4	32	5
7	93	5	23	5
8	95	4	24	4
9	134	6	34	5
10	114	6	29	5
11	120	7	30	4
12	150	6	38	4

### 3.3 Proses Clustering

Vektor Fitur yang telah dibuat diolah menggunakan K-Means Clustering dengan nilai K yang sudah ditentukan, untuk membentuk kelompok data yang memiliki kemiripan. Pada penelitian ini penulis menggunakan Matlab 2017a sebagai aplikasi bantuan, dengan fungsi bawaan dari Matlab yaitu *kmeans*.

### 3.4 Proses Peramalan

Setelah data dikelompokkan menggunakan K-Means Clustering, berikutnya untuk setiap kelompok dilakukan proses pembentukan jaringan Backpropagation. Pada proses pelatihan jaringan Backpropagation, masukan yang digunakan adalah Vektor Fitur kolom  $X_1, X_2, \dots, X_{M-1}$  dan Targetnya adalah  $X_M$ . Misal jika  $M = 5$ , maka masukannya adalah  $X_1, X_2, X_3, X_4$  dan targetnya adalah  $X_5$ .

Berdasarkan *trial and error* diperoleh bahwa komposisi pembagian data yang digunakan untuk membangun jaringan Backpropagation, yaitu 80% untuk data latih dan 20% untuk data uji.

Pada penelitian ini Matlab 2017a digunakan sebagai alat bantu, dimana untuk membangun jaringan Backpropagation digunakan fungsi bawaan dari matlab yaitu *feedforwardnet*, dengan parameter banyaknya hidden layer yang akan digunakan untuk proses pelatihan. Fungsi *train* untuk pelatihan jaringan yang telah diinisiasi, yaitu data latih. Fungsi *sim* untuk mendapatkan keluaran dari masukan yang diberikan. Gambar 4 adalah jendela Neural Network Training hasil dari pelatihan salah satu eksperimen. Setelah proses pelatihan jaringan selesai, berikutnya adalah proses testing menggunakan data uji. Domain keluaran dari proses testing adalah  $[0, 1]$ , supaya keluran proses testing dapat digunakan sebagai hasil peramalan perlu dilakukan transformasi data menggunakan persamaan 14, tahap ini disebut sebagai denormalisasi. Setelah hasil peramalan didapat berikutnya adalah tahap evaluasi hasil peramalan untuk menentukan kualitas peramalan menggunakan metode Mean Absolute Deviation (MAD).

### 3.5 Pengujian

Pada penelitian ini dilakukan beberapa pengujian, yaitu :

#### 3.5.1 Pengujian pengelompokan menggunakan K-Means Clustering dengan Metode *Silhouette*

Metode *Silhouette* menggabungkan dua metode yaitu metode *cohesion* yang berfungsi untuk mengukur seberapa dekat relasi antara objek dalam sebuah *cluster*, dan metode *separation* yang berfungsi untuk mengukur

seberapa jauh sebuah cluster terpisah dengan *cluster* lain.

Pada penelitian ini data yang akan dikelompokkan adalah Vektor Fitur yang merupakan data penjualan yang disusun dengan pola terurut, dimana panjang Vektor (M) ditentukan secara acak. Banyak kelompok (K) pada K-Means ditentukan secara acak, namun memiliki batasan yang ditentukan dengan formula Sturgess.

Pengujian ini bertujuan untuk mendapatkan nilai M dan K yang tepat pada proses pengelompokan menggunakan K-Means. Domain nilai M yang digunakan adalah 14 hingga 18, dan domain nilai K yang digunakan adalah 2 hingga 12 (hasil dari formula Sturgess). Hasil dari pengujian ini seperti pada tabel 4. Dari pengujian ini didapat nilai Silhouette yang terbaik adalah 0.0349, yang diperoleh dari pengelompokan yang menggunakan nilai M dan K-nya 16 dan 12.

### 3.5.2 Pengujian proses peramalan menggunakan Backpropagation dengan Metode MAD

Proses peramalan pada penelitian ini menggunakan metode Backpropagation, dengan data masukannya adalah Vektor Fitur hasil pengelompokan K-Means yang panjang vektornya (M) adalah 16 dan banyak kelompoknya (K) adalah 12. Atribut Vektor Fitur dibagi menjadi 2, yaitu input dan target, dimana inputnya adalah  $X_1, X_2, \dots, X_{15}$  dan targetnya adalah  $X_{16}$ . Data pada setiap kelompok akan dibagi menjadi 2 bagian, bagian pelatihan dan bagian ujicoba, dengan komposisi 80% dan 20%, dimana banyak datanya adalah 1812.

Proses pengujian ini, diawali dengan proses pembentukan jaringan pada setiap kelompok hasil dari K-Means clustering, kemudian proses Ujicoba. Penentuan nilai MAD-nya dilakukan dengan denormalisasi output dari jaringan Backpropagation, kemudian membandingkannya dengan target.

Pada hasil pengujian ini (tabel 5) didapatkan rata-rata nilai MAD proses pelatihan adalah 5.5 dan untuk proses pengujian adalah 4.5.

### 3.5.3 Perbandingan model peramalan menggunakan Backpropagation dengan model peramalan K-Means dan Backpropagation

Pada penelitian sebelumnya sistem peramalan hanya menggunakan Backpropagation dan kualitas peramalan yang didapat belum memuaskan. Pada penelitian

ini sistem peramalan dimodifikasi, sehingga sebelum dilakukan peramalan menggunakan Backpropagation, data dikelompokkan terlebih dahulu sehingga data awal yang heterogen dapat dijadikan lebih homogen. Tujuan perbandingan ini adalah untuk mengetahui apakah modifikasi sistem peramalan dengan mengintegrasikan K-Means Clustering dan Backpropagation lebih baik dari sistem peramalan yang hanya menggunakan Backpropagation saja.

Pada proses peramalan yang hanya menggunakan Backpropagation saja didapat nilai MAD-nya sebesar 6 untuk pelatihan dan 5,5 untuk ujicoba, sedangkan modifikasi sistem peramalan dengan mengintegrasikan K-Means Clustering dan Backpropagation sebesar 5.5 untuk pelatihan dan 4.5 untuk ujicoba. Hal ini membuktikan bahwa modifikasi sistem peramalan yang dilakukan telah berhasil memperbaiki kualitas peramalan sistem yang sebelumnya.

## 4. SIMPULAN

Berdasarkan hasil penelitian dan pembahasan tentang Integrasi Metode K-Means Clustering dan Backpropagation pada Model Peramalan Penjualan dapat ditarik kesimpulan sebagai berikut :

1. Integrasi metode K-Means *Clustering* dan *Backpropagation* pada model peramalan penjualan mampu memperbaiki kualitas peramalan dari sistem peramalan sebelumnya yang hanya menggunakan metode *Backpropagation* saja.
2. Pada proses pengelompokan didapat kombinasi panjang Vektor Fitur (M) dan banyak kelompok (K) yang terbaik adalah 16 dan 12, dimana nilai *Silhouette Coefficient* yang didapat adalah 0.0349.
3. Pada proses peramalan didapat nilai rata-rata Mean Absolute Deviation (MAD) dari proses pelatihan dan proses ujicoba adalah 5.5 dan 4.5. Nilai MAD ini lebih baik dari pada sistem yang hanya menggunakan Backpropagation untuk sistem peramalannya, yang memiliki nilai MAD sebesar 6 dan 5.5 untuk proses pelatihan dan proses ujicobanya.

## 5. SARAN

Penelitian Integrasi Metode K-Means Clustering dan Backpropagation pada Model Peramalan Penjualan hanya menggunakan 1 kali percobaan pada arsitektur jaringan Backpropagation (proses peramalan), sehingga untuk penelitian selanjutnya dapat menggunakan lebih banyak percobaan agar dapat memperoleh model yang lebih baik.

## DAFTAR PUSTAKA

- [1] L. Fausett, *Fundamentals of Neural Networks: Architectures, Algorithms And Applications*. New Jersey: Prentice-Hall, 1994.
- [2] J. J. Siang, *Jaringan syaraf tiruan dan pemrogramannya menggunakan matlab*. Yogyakarta : Andi, 2005.
- [3] Solikhin, R. Rachmatullah, and E. Riyanto, "PENGEMBANGAN SISTEM INFORMASI REGISTRASI SEMINAR, WORKSHOP DAN PELATIHAN MENGGUNAKAN METODE SYSTEM DEVELOPMENT LIFE CYCLE MODEL WATERFALL (STUDI KASUS STMIK HIMSYA SEMARANG)," vol. Vol 10, No, 2014. Available: <https://ojs.stmik-himsya.ac.id/index.php/JHT/article/view/59>.
- [4] A. G. Salman and Y. L. Prasetio, "IMPLEMENTASI JARINGAN SYARAF TIRUAN RECURRENT DENGAN METODE PEMBELAJARAN GRADIENT DESCENT ADAPTIVE LEARNING RATE UNTUK PENDUGAAN CURAH HUJAN BERDASARKAN PEUBAH ENSO," 2010.
- [5] A. Hikmah, A. Agoestanto, and R. Arifudin, "PERAMALAN DERET WAKTU DENGAN MENGGUNAKAN AUTOREGRESSIVE (AR), JARINGAN SYARAF TIRUAN RADIAL BASIS FUNCTION (RBF) DAN HIBRID AR-RBF PADA INFLASI INDONESIA," *UNNES J. Math.*, vol. Vol. 7(2);, 2017. Available: <https://journal.unnes.ac.id/sju/index.php/ujm>.
- [6] B. Suhada, "PERAMALAN PRODUKSI GULA NASIONAL MELALUI PENDEKATAN ARTIFICIAL NEURAL NETWORK," *J. Deriv.*, vol. Vol. 3(1);, 2009.
- [7] A. F. Achmalia and S. Walid, "PERAMALAN PENJUALAN SEMEN MENGGUNAKAN BACKPROPAGATION NEURAL NETWORK DAN RECURRENT NEURAL NETWORK," *UNNES J. Math.*, 2020.
- [8] N. Kohzadi, M. S. Boyd, B. Kermanshahi, and I. Kaastra, "A comparison of artificial neural network and time series models for forecasting commodity prices," *Neurocomputing*, vol. Vol. 10: 1, 1996.
- [9] G. Zhang and M. Y. Hu, "Neural network forecasting of the British Pound/US Dollar exchange rate," *Omega*, vol. Vol. 26(4), 1998.
- [10] A. S. Rachman, I. Cholissodin, and M. A. Fauzi, "Peramalan Produksi Gula Menggunakan Metode Jaringan Syaraf Tiruan Backpropagation pada PG Candi Baru Sidoharjo," vol. Vol. 2, No, 2018. Available: <https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/1324/471>.
- [11] R. Diputra, M., Dewi, C., & Wihandika, "Prediksi Tingkat Keuntungan Usaha Peternakan Itik Alabio Petelur Menggunakan Jaringan Syaraf Tiruan Backpropagation (Kasus Di Kabupaten Hulu Sungai Utara Kalimantan Selatan)," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no, 2017. Available: <https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/959>.
- [12] M. I. Ghozali, "Rantai Pasok Beras Pada Bulog Berbasis Neural Network," vol. Vol 7, No, 2016. Available: <https://jurnal.umk.ac.id/index.php/simet/article/view/790>.
- [13] Y. Darmi and A. Setiawan, "PENERAPAN METODE CLUSTERING K-MEANS DALAM PENGELOMPOKAN PENJUALAN PRODUK," *J. Media Infotama*, vol. Vol.12No.2, 2016. Available: <https://docplayer.info/69618221-Penerapan-metode-clustering-k-means-dalam-pengelompokan-penjualan-produk.html>.
- [14] F. Nasari and S. Darma, "PENERAPAN K-MEANS CLUSTERING PADA DATA PENERIMAAN MAHASISWA BARU," vol. Vol 3, No, 2015. Available: <https://ojs.amikom.ac.id/index.php/semnasteknomedia/article/view/837/801>.
- [15] M. Anike, Suyoto, and Ernawati, "Pengembangan Jaringan Syaraf Tiruan dalam Memprediksi Jumlah Dokter Keluarga Menggunakan Backpropagation (Studi Kasus : Regional X Cabang Palu)," 2012.