



## **Analisis Pengaruh Fungsi Aktivasi, Learning Rate Dan Momentum Dalam Menentukan Mean Square Error (MSE) Pada Jaringan Saraf Restricted Boltzmann Machines (RBM)**

### ***Analysis of The Influence Activation Function, Learning Rate And Momentum in Determining Mean Square Error (MSE) in Restricted Boltzmann Machines (RBM) Neural Network***

Susilawati <sup>1)</sup>\* & Muhathir <sup>1)</sup>

1) Program Studi Teknik Informatika, Fakultas Teknik Universitas Medan Area, Indonesia

\*Corresponding Email: [susi.shilawati@gmail.com](mailto:susi.shilawati@gmail.com)

---

#### **Abstrak**

Restricted boltzmann machines (RBM) merupakan algoritma pembelajaran jaringan syaraf tanpa pengawaas (*unsupervised learning*) yang hanya terdiri dari dua lapisan yang *visible layer* dan *hidden layer*. Kinerja RBM sangat dipengaruhi oleh parameter-parameternya seperti fungsi aktivasi yang digunakan untuk mengaktifkan neuron pada jaringan dan *learning rate* serta *momentum* untuk mempercepat proses pembelajaran. Pemilihan fungsi aktivasi yang tepat sangat mempengaruhi kinerja dalam menentukan *Mean Square Error* (MSE) pada jaringan saraf RBM. Fungsi aktivasi yang digunakan pada jaringan RBM adalah fungsi aktivasi sigmoid. Beberapa varian dari fungsi aktivasi sigmoid seperti fungsi sigmoid biner dan sigmoid tangen hiperbolik (*tanh*). Dengan menggunakan dataset MNIST untuk pembelajaran dan pengujian, terlihat bahwa tingkat keberhasilan untuk klasifikasi pada fungsi aktivasi sigmoid biner, ditentukan oleh nilai MSE yang kecil. Berbeda dengan fungsi aktivasi tangen nilai MSE menaik seiring bertambahnya jumlah epoch. Fungsi aktivasi sigmoid biner dengan learning rate 0.05 dan momentum 0.7 memiliki tingkat pengenalan tulisan tangan yang tinggi sebesar 93.42%, diikuti dengan learning rate 0.01 momentum 0.9 yakni 91.92%, learning rate 0.05 momentum 0.5 yakni 91.31%, learning rate 0.01 momentum 0.7 sebesar 90.56% dan terakhir learning rate 0.01 momentum 0.5 sebesar 87.49%.

**Kata Kunci:** Jaringan Saraf, Mean Square Error (MSE), Restricted boltzmann machines (RBM).

#### **Abstract**

*Restricted Boltzmann Machines (RBM) is a neural network unsupervised learning algorithm which only consists of two layers of the visible layer and the hidden layer. RBM performance is strongly influenced by parameters such as the activation function that is used to activate neurons in the network and the learning rate and momentum to accelerate the learning process. Selection of the activation function corectly very influence the performance in determining the Mean Square Error (MSE) on RBM neural network. Activation function that is used on RBM network is the sigmoid activation function. Several variants of the sigmoid activation function like binary sigmoid and sigmoid hyperbolic tangent (tanh). By using datasets MNIST for learning and testing, it appears that the success rate for the classification of the binary sigmoid activation function, is determined by the value of MSE is small. Unlike the tangent activation function MSE ascending of rising number of epoch. Activation function binary sigmoid with the learning rate of 0.05 and momentum 0.7 has a recognition rate of handwriting a high namely 93.42%, followed by the learning rate 0.01 momentum 0.9 namely 91.92%, for learning rate 0.05 momentum 0.5 namely 91.31%, learning rate 0.01 and momentum 0.7 is 90.56% and the last learning rate 0.01 and momentum 0.5 namely 87.49%.*

**Keywords:** Neural Network, Mean Square Error (MSE), Restricted boltzmann machines (RBM).

How to Cite: Susilawati, & Muhathir. (2019). Analisis Pengaruh Fungsi Aktivasi, Learning Rate Dan Momentum Dalam Menentukan Mean Square Error (MSE) Pada Jaringan Saraf Restricted Boltzmann Machines (RBM). *JITE(Journal of Informatics and Telecommunication Engineering)*. 2 (2):77-91

---

## PENDAHULUAN

Penelitian tentang jaringan saraf tiruan atau *artificial neural network* (ANN) dengan menggunakan algoritma Restricted Boltzmann Machines (RBM) (Smolensky, 1986) telah banyak dilakukan. Kemampuan dalam meningkatkan pengolahan pada chip seperti GPU, meningkatkan kecepatan proses pembelajaran RBM untuk ukuran data yang sangat besar.

RBM merupakan metode pembelajaran jaringan saraf tiruan tanpa pengawasan (*unsupervised learning*) dengan menggunakan arsitektur jaringan berulang (*recurrent network*) yang hanya terdiri dari dua lapisan yakni lapisan terlihat (*visible layer*) dan lapisan tersembunyi (*hidden layer*). RBM telah diterapkan untuk penyelesaian berbagai masalah seperti klasifikasi (Ahmed et al, 2008; Bengio et al, 2007; Larochelle et al, 2007; Ranzato et al, 2007; Lee et al, 2009), pemodelan tekstur (Osindero et al, 2008), pemodelan gerak (Taylor et al, 2007; Taylor et al, 2009), pengambilan informasi (Ranzato et al 2008; Salakhudinov et al 2007; Torralba et al, 2008), robotika (Hadsell et al, 2008), dan pengenalan pola seperti pengenalan tulisan tangan angka (Susilawati, 2017). Dari hasil penelitian tersebut diperoleh kesimpulan awal bahwa algoritma RBM dapat

diimplementasikan pada berbagai masalah.

Fungsi aktivasi sangat berperan dalam mengaktifkan setiap neuron pada jaringan saraf tiruan serta menentukan keluaran dari suatu jaringan saraf tiruan. Pilihan fungsi aktivasi yang dapat digunakan dalam metode RBM seperti fungsi sigmoid biner dan tangen hiperbolik (*tanh*). Pemilihan fungsi aktivasi yang tepat tentunya akan mempengaruhi kinerja jaringan RBM, seperti dalam hal kecepatan pemrosesan data dan tingkat keakuratan hasil yang diperoleh (Hinton, 2002).

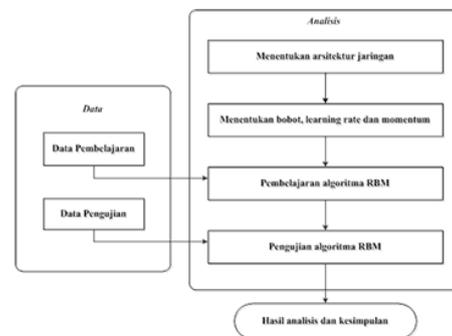
Pertimbangan yang harus diperhatikan saat menggunakan jaringan RBM adalah kemungkinan terjadinya konvergen pada lokal minimum dan penentuan kriteria penghentian pembelajaran agar waktu tercapainya konvergen jaringan tidak terlalu lama. Hal ini diawali dengan pengaturan bobot hingga pembaharuan nilai bobot agar diperoleh bobot-bobot yang baik selama proses pembelajaran. Setiap perubahan bobot yang terjadi selama siklus pola pembelajaran diharapkan dapat mengurangi besarnya nilai kesalahan (*mean square error*), sampai diperoleh tingkat kesalahan yang paling minimum dengan bobot yang stabil (Hinton, 2002).

Untuk mengatasi permasalahan tersebut diperlukan parameter momentum yang berfungsi untuk menstabilkan proses pembelajaran dan mencegah sistem dari konvergen pada lokal minimum, serta parameter learning rate untuk mempercepat laju pembelajaran. Selain momentum, nilai learning rate memiliki peranan pada jaringan RBM. Menurut Hinton (2010), jaringan RBM yang memiliki nilai learning rate terlalu besar dapat menyebabkan terjadinya peningkatan nilai bobot secara drastis, sehingga menyebabkan terjadinya kesalahan pada saat proses rekonstruksi. Nilai learning rate yang rendah dapat menurunkan tingkat kesalahan pada saat proses rekonstruksi. Tidak adanya ketentuan yang pasti dalam menentukan nilai parameter-parameter tersebut saat proses pembelajaran jaringan RBM, menyebabkan dibutuhkan pengalaman praktis agar diperoleh hasil maksimum sesuai dengan pola data yang digunakan (Hinton, 2002).

Dalam penelitian ini, dilakukan analisa tentang penentuan fungsi aktivasi, pemberian nilai momentum dan nilai learning rate yang tepat untuk menentukan mean square error pada jaringan RBM dalam pemrosesan data.

## METODE PENELITIAN

Tahapan penelitian yang dilakukan adalah seperti gambar 1.



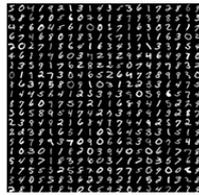
Gambar 1. Tahapan penelitian

### 2.2.1 Data Yang Digunakan

Data yang digunakan dalam penelitian ini adalah dataset MNIST (Yann Le Cun, 2010). Dataset MNIST dibangun oleh Institute National Standard Teknologi (NIST). Dataset MNIST terdiri dari gambar tulisan tangan angka 0 sampai 9 dan dibagi dalam 60.000 contoh data pembelajaran dan 10.000 data untuk pengujian. Berikut adalah tabel distribusi data pelatihan dan pengujian dataset MNIST.

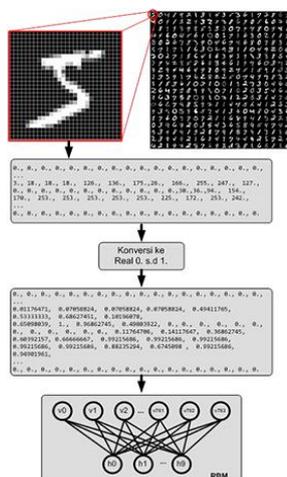
Gambar 2 berikut merupakan visualisasi dataset MNIST sebanyak 400 sample data. Dimana, data tersebut terdiri dari angka 0 sampai 9 dengan karakteristik yang berbeda-beda, ada data yang tampak lebih tebal dan ada yang tampak lebih tipis selain itu ada pula data

yang tampak dengan jelas dan ada yang tampak tidak jelas.



Gambar 2. 400 sampel dataset

Setiap pixel gambar diwakili oleh nilai antara 0 sampai 255, di mana 0 adalah hitam, 255 adalah putih, serta warna keabuan dari piksel. Ukuran gambar terpusat pada 28 x 28 piksel. Dimana pembacaan piksel gambar dilakukan secara perbaris. Data pada masing-masing piksel selanjutnya dibagi dengan angka 255 untuk memperoleh bilangan real antara 0 sampai 1. Data real ini selanjutnya dijadikan data masukan pada pembelajaran RBM yang diwakilkan pada lapisan terlihat (*visible layer*), kemudian digunakan untuk menginisialisasi nilai pada node yang ada pada lapisan tersembunyi (*hidden layer*). Adapun teknik pembacaan data tersebut dapat dilihat pada gambar 3



Gambar 3. Teknik pembacaan dataset

### 2.2.2 Algoritma RBM

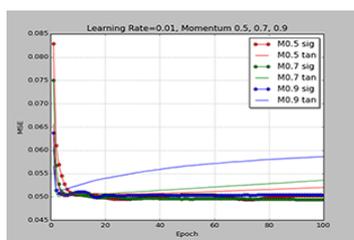
Proses pembelajaran RBM merupakan pembelajaran tanpa pengawasan (*unsupervised learning*), artinya tidak ada target data saat proses pembelajaran dilakukan. Hasil pembelajaran akan berkelompok sesuai klasifikasi masing-masing data. Output aktual yang dihasilkan pada awal proses akan digunakan sebagai input pada proses berikutnya. Kesalahan (*error*) pada output jaringan diperoleh dari selisih antara output pada fase positif dan output pada fase negatif. Epoch diperlukan sebagai batasan untuk proses pembelajaran.

Proses pembelajaran RBM dimulai dari beberapa tahap yaitu :

#### a. Penentuan Arsitektur Jaringan

Secara umum bentuk arsitektur jaringan saraf yang digunakan dapat dilihat pada gambar 4 Arsitektur jaringan disesuaikan dengan jaringan RBM yang terdiri dari dua lapisan yaitu lapisan unit terlihat (*visible*) dan lapisan unit tersembunyi (*hidden*). Jumlah node pada lapisan terlihat akan disesuaikan dengan jumlah piksel data yang akan digunakan sebagai data masukan, sedangkan jumlah node pada lapisan tersembunyi akan disesuaikan dengan jumlah pengelompokan data.

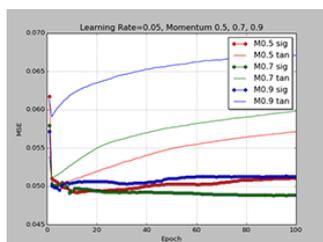




Gambar 6. Paduan Fungsi Aktivasi *Tanh* dan Sigmoid, *Learning Rate* 0.01, *Momentum* 0.5, 0.7, 0.9

Gambar 6 menampilkan grafik nilai MSE yang diperoleh dari fungsi aktivasi sigmoid dan tanh. Secara ringkas dapat dilihat bahwa grafik merah, hijau dan biru yang tebal untuk fungsi aktivasi sigmoid tampak cenderung mendatar dengan selisih nilai yang kecil, sedangkan grafik merah, hijau dan biru yang tipis untuk fungsi aktivasi tanh, tampak nilai MSE yang terus menaik secara perlahan hingga epoch terakhir.

Perbedaan nilai MSE yang dihasilkan menggunakan fungsi aktivasi sigmoid dan tanh dengan learning rate 0.05, momentum 0.5, 0.7 dan 0.9 lebih jelas dapat dilihat pada gambar 7 berikut:

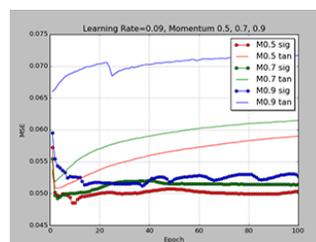


Gambar 7. Paduan Fungsi Aktivasi *Tanh* dan Sigmoid, *Learning Rate* 0.05, *Momentum* 0.5, 0.7, 0.9

Gambar 7 menampilkan grafik nilai MSE yang diperoleh dari fungsi aktivasi sigmoid dan tanh. Secara ringkas dapat dilihat bahwa untuk aktivasi sigmoid, grafik merah dan biru tampak nilai MSE

naik turun pada epoch tertentu, sedangkan grafik hijau tampak cenderung menurun secara perlahan dengan selisih nilai yang kecil. Untuk aktivasi tanh nilai MSE cenderung menaik secara perlahan hingga epoch terakhir.

Perbedaan nilai MSE yang dihasilkan menggunakan fungsi aktivasi sigmoid dan tanh dengan learning rate 0.09, momentum 0.5, 0.7 dan 0.9 lebih jelas dapat dilihat pada gambar 8 berikut:



Gambar 8. Paduan Fungsi Aktivasi *Tanh* dan Sigmoid, *Learning Rate* 0.09, *Momentum* 0.5, 0.7, 0.9

Gambar 8 menampilkan grafik nilai MSE yang diperoleh dari fungsi aktivasi sigmoid dan tanh tidak stabil yakni cenderung menaik hingga akhir epoch.

Tabel 1 berikut adalah ringkasan nilai MSE terkecil yang dihasilkan berdasarkan proses pembelajaran yang dilakukan.

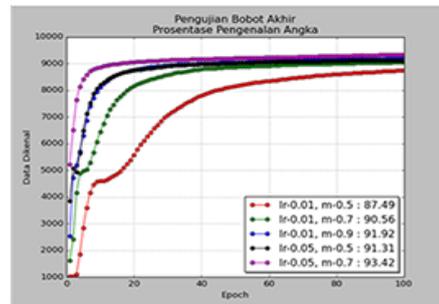
Tabel 1. Hasil MSE terkecil dari tahapan pembelajaran

Fungsi Aktivasi	Pembelajaran Tahap Ke	Learning Rate (ε)	Momentum (m)	MSE	Epoch ke
Sigmoid	I	0.01	0.5	0.0493	82
			0.7	0.0494	95
			0.9	0.0499	26
	II	0.05	0.5	0.0491	12
			0.7	0.0488	88
			0.9	0.0495	5
	III	0.09	0.5	0.0485	8
			0.7	0.0492	3
			0.9	0.0512	47
Tanh	IV	0.01	0.5	0.0501	5
			0.7	0.0499	3
			0.9	0.0514	3
	V	0.05	0.5	0.0504	5
			0.7	0.0511	2
			0.9	0.0590	2
	VI	0.09	0.5	0.0508	2
			0.7	0.0518	2
			0.9	0.0689	1

Kriteria penentuan nilai parameter learning rate dan momentum yang akan digunakan untuk tahap pengujian, dilihat berdasarkan grafik MSE yang dihasilkan dan hasil visualisasi pengelompokan dataset pada proses pembelajaran. Parameter dengan fungsi aktivasi tanh tidak terlihat memenuhi persyaratan sama sekali. Sedangkan beberapa parameter *learning rate* dan *momentum* yang menggunakan fungsi aktivasi sigmoid memenuhi persyaratan. Parameter *learning rate* dan *momentum* yang memiliki persyaratan tersebut adalah learning rate 0.01 dengan momentum 0.5, 0.7, 0.9 dan learning rate 0.05 dengan momentum 0.5, 0.7.

Visualisasi klasifikasi dataset dari pilihan nilai-nilai parameter tersebut menunjukkan hasil yang baik dan jelas. Nilai-nilai dari parameter ini akan digunakan untuk proses pengujian terhadap data pengujian MNIST untuk mengetahui tingkat persentase pengenalan data pengujian MNIST. Dari hasil proses tahap pembelajaran diperoleh beberapa bobot akhir yang baik untuk digunakan pada tahap pengujian. Bobot akhir yang baik dinilai dari nilai MSE dan hasil klasifikasi yang baik.

Hasil pengujian bobot akhir dan persentase pengenalan data MNIST dapat dilihat pada gambar 9.



Gambar 9. Grafik persentase pengenalan dataset pengujian

Gambar 9 menunjukkan bahwa parameter *learning rate* 0.01 dengan *momentum* 0.5 berwarna merah dapat mengenali dataset pengujian tulisan tangan angka maksimum 87.49%. Dilanjutkan dengan learning rate 0.01 dengan momentum 0.7 berwarna hijau yang dapat mengenal dataset pengujian maksimum sebesar 90.56%. Berikutnya learning rate 0.01 dengan momentum 0.9 berwarna biru dan learning rate 0.05 dengan momentum 0.5 berwarna hitam, memiliki kinerja yang hampir sama, dapat mengenali dataset masing-masing 91.92% dan 91.31%. Terakhir, parameter learning rate 0.05 dengan momentum 0.7 memiliki kinerja yang paling tinggi, mampu mengenali dataset pengujian sebesar 93.42%.

Dari hasil pengujian terlihat bahwa pemakaian fungsi aktivasi sigmoid dengan learning rate 0.05 dan momentum 0.7

memiliki tingkat pengenalan dataset angka tulisan tangan yang tinggi sebesar 93.42%. Diikuti dengan pilihan nilai learning rate 0.01 dan momentum 0.9 memiliki tingkat pengenalan dataset angka tulisan tangan sebesar 91.92%. Selanjutnya pilihan nilai learning rate 0.05 dan nilai momentum 0.5 diperoleh tingkat pengenalan dataset angka tulisan tangan sebesar 91.31%, pilihan nilai learning rate 0.01 dan momentum 0.7 diperoleh tingkat pengenalan sebesar 90.56% dan terakhir nilai learning rate 0.01 dan momentum 0.5 diperoleh tingkat pengenalan sebesar 87.49%.

### 3.2 Pembahasan

#### 3.2.1 Pembelajaran

Pembelajaran dilakukan dengan membangun jaringan RBM yang terdiri dari 784 neuron *visible layer* dan 10 neuron pada *hidden layer*. Pembacaan dataset dilakukan secara bertahap, batch demi batch, mengingat jumlah data pelatihan yang besar yakni 60000 data tulisan tangan. Batch digunakan untuk mempermudah pembacaan data yang besar. Jumlah 60000 data pelatihan dibagi dengan jumlah batch yang dipilih yakni 64 data untuk satu batch. Jadi 60000 data pelatihan dibagi dengan 64 menjadi 937 jumlah batch yang akan diproses untuk setiap epoch. Hasil MSE dan bobot-bobot yang diperoleh dicatat dan disimpan untuk

digunakan sebagai parameter pada tahap pengujian.

Untuk tahap pembelajaran, parameter yang akan diuji adalah *learning rate* 0.01, 0.05, 0.09 dan *momentum* 0.5, 0.7, 0.9, dengan fungsi aktivasi *sigmoid* dan *tanh* serta pembatasan epoch sebanyak 100 epoch.

#### a. Pembelajaran Tahap I dengan Fungsi Aktivasi Sigmoid

Pada pembelajaran tahap pertama ini fungsi aktivasi yang digunakan adalah sigmoid biner dengan nilai parameter *learning rate* yang dipilih adalah 0.01 dengan nilai *momentum* bervariasi 0.5, 0.7 dan 0.9, jumlah *epoch* maksimum 100 *epoch*. Hasil lengkap nilai MSE yang diperoleh dapat dilihat pada tabel 2.

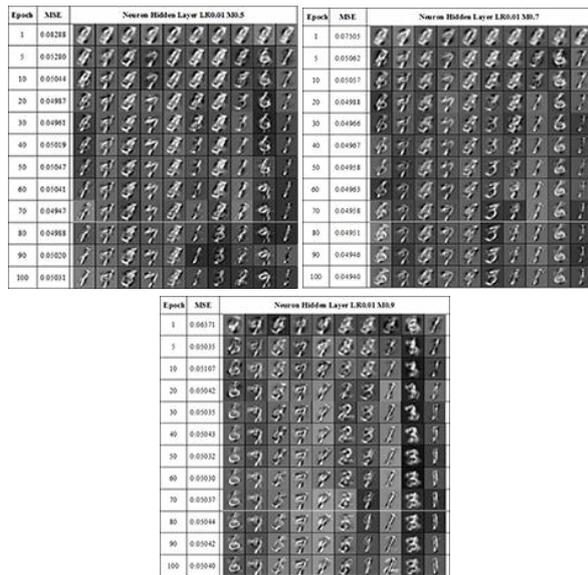
Tabel 2. Nilai MSE Fungsi Aktivasi Sigmoid dengan Learning Rate 0.01, Momentum 0.5, 0.7, 0.9

Ep	0	1	2	3	4	5	6	7	8	9	Mom
0	0.0828	0.0610	0.0569	0.0543	0.0528	0.0516	0.0509	0.0506	0.0504	0.0504	0.5
	0.0750	0.0567	0.0528	0.0511	0.0506	0.0505	0.0505	0.0505	0.0505	0.0505	0.9
1	0.0637	0.0513	0.0507	0.0505	0.0505	0.0505	0.0505	0.0507	0.0509	0.0510	0.9
	0.0504	0.0503	0.0503	0.0503	0.0503	0.0502	0.0502	0.0501	0.0500	0.0498	0.5
2	0.0504	0.0503	0.0502	0.0501	0.0500	0.0499	0.0498	0.0498	0.0498	0.0498	0.7
	0.0511	0.0510	0.0509	0.0507	0.0505	0.0499	0.0499	0.0500	0.0503	0.0504	0.9
3	0.0497	0.0496	0.0495	0.0495	0.0494	0.0494	0.0494	0.0495	0.0495	0.0496	0.5
	0.0499	0.0499	0.0499	0.0499	0.0498	0.0497	0.0497	0.0496	0.0496	0.0496	0.7
4	0.0504	0.0503	0.0503	0.0503	0.0502	0.0503	0.0503	0.0503	0.0503	0.0503	0.9
	0.0496	0.0496	0.0496	0.0496	0.0496	0.0496	0.0495	0.0495	0.0495	0.0495	0.7
5	0.0504	0.0503	0.0503	0.0503	0.0503	0.0503	0.0502	0.0503	0.0503	0.0503	0.9
	0.0495	0.0495	0.0495	0.0495	0.0496	0.0496	0.0496	0.0496	0.0496	0.0496	0.7
6	0.0505	0.0505	0.0505	0.0504	0.0505	0.0505	0.0505	0.0505	0.0505	0.0505	0.9
	0.0502	0.0502	0.0502	0.0500	0.0499	0.0498	0.0497	0.0496	0.0495	0.0494	0.5
7	0.0496	0.0496	0.0496	0.0496	0.0496	0.0496	0.0496	0.0496	0.0495	0.0495	0.7
	0.0505	0.0505	0.0505	0.0505	0.0505	0.0505	0.0505	0.0505	0.0505	0.0505	0.9
8	0.0494	0.0493	0.0494	0.0494	0.0495	0.0495	0.0495	0.0495	0.0495	0.0495	0.7
	0.0504	0.0504	0.0504	0.0504	0.0504	0.0504	0.0504	0.0504	0.0504	0.0504	0.9
9	0.0495	0.0495	0.0495	0.0495	0.0495	0.0495	0.0495	0.0495	0.0495	0.0495	0.7
	0.0502	0.0502	0.0502	0.0502	0.0502	0.0502	0.0502	0.0502	0.0502	0.0502	0.5
9	0.0494	0.0494	0.0494	0.0494	0.0494	0.0494	0.0494	0.0494	0.0494	0.0494	0.7
	0.0504	0.0504	0.0504	0.0504	0.0504	0.0504	0.0504	0.0504	0.0504	0.0504	0.9

Dari tabel 2 terlihat bahwa untuk *learning rate* 0.01 dengan *momentum* 0.5 nilai MSE terkecil diperoleh pada *epoch* ke-

72 sebesar 0.0493. Diikuti dengan momentum 0.7 pada epoch ke-95 sebesar 0.0494 dan momentum 0.9 pada epoch ke-26 sebesar 0.0499.

Gambar 10 berikut merupakan visualisasi klasifikasi data yang dihasilkan



Gambar 10. Visualisasi klasifikasi data Fungsi Aktivasi Sigmoid dengan Learning Rate 0.01, Momentum 0.5, 0.7, 0.9

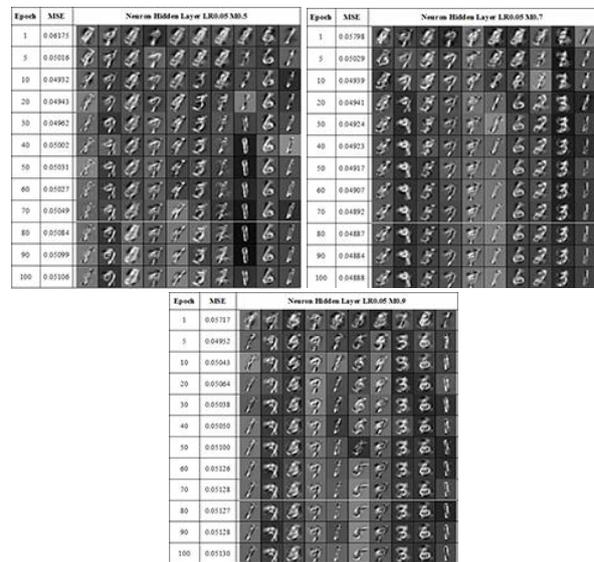
**b. Pembelajaran Tahap II dengan Fungsi Aktivasi Sigmoid**

Pada pembelajaran tahap kedua ini fungsi aktivasi yang digunakan adalah sigmoid biner dengan nilai parameter learning rate yang dipilih adalah 0.05 dengan nilai momentum bervariasi 0.5, 0.7 dan 0.9, jumlah epoch maksimum 100 epoch. Hasil lengkap nilai MSE yang diperoleh dapat dilihat pada tabel 3.

Tabel 3 Nilai MSE Fungsi Aktivasi Sigmoid Biner Learning Rate 0.05, Momentum 0.5, 0.7, 0.9

Ep.	0	1	2	3	4	5	6	7	8	9	Mom.
0	0.0617	0.0510	0.0507	0.0505	0.0501	0.0499	0.0498	0.0497	0.0495	0.0493	0.5
	0.0579	0.0503	0.0500	0.0500	0.0502	0.0504	0.0503	0.0499	0.0491	0.0493	0.7
1	0.0571	0.0500	0.0498	0.0497	<b>0.0495</b>	0.0500	0.0503	0.0504	0.0503	0.0504	0.9
	0.0492	<b>0.0491</b>	0.0492	0.0493	0.0495	0.0494	0.0494	0.0494	0.0494	0.0494	0.5
2	0.0496	0.0497	0.0497	0.0496	0.0496	0.0496	0.0496	0.0496	0.0495	0.0494	0.7
	0.0505	0.0506	0.0506	0.0506	0.0506	0.0506	0.0506	0.0506	0.0506	0.0506	0.9
3	0.0494	0.0494	0.0494	0.0494	0.0495	0.0495	0.0495	0.0495	0.0495	0.0495	0.5
	0.0493	0.0493	0.0492	0.0492	0.0492	0.0492	0.0492	0.0492	0.0492	0.0492	0.7
4	0.0506	0.0506	0.0506	0.0506	0.0506	0.0506	0.0505	0.0504	0.0503	0.0503	0.9
	0.0496	0.0496	0.0496	0.0497	0.0497	0.0498	0.0498	0.0499	0.0499	0.0500	0.5
5	0.0492	0.0492	0.0492	0.0492	0.0492	0.0492	0.0492	0.0492	0.0492	0.0492	0.7
	0.0503	0.0503	0.0503	0.0503	0.0504	0.0504	0.0504	0.0504	0.0504	0.0505	0.9
6	0.0500	0.0501	0.0501	0.0501	0.0502	0.0502	0.0502	0.0502	0.0502	0.0503	0.5
	0.0492	0.0492	0.0492	0.0491	0.0492	0.0492	0.0491	0.0492	0.0491	0.0491	0.7
7	0.0505	0.0505	0.0505	0.0506	0.0506	0.0507	0.0508	0.0509	0.0510	0.0510	0.9
	0.0510	0.0510	0.0511	0.0511	0.0511	0.0512	0.0512	0.0512	0.0512	0.0512	0.5
8	0.0502	0.0502	0.0502	0.0502	0.0503	0.0503	0.0503	0.0503	0.0503	0.0504	0.7
	0.0490	0.0490	0.0490	0.0489	0.0489	0.0489	0.0489	0.0489	0.0489	0.0489	0.9
9	0.0512	0.0513	0.0512	0.0512	0.0512	0.0512	0.0512	0.0512	0.0512	0.0512	0.5
	0.0505	0.0505	0.0505	0.0506	0.0506	0.0507	0.0507	0.0508	0.0508	0.0508	0.7
10	0.0489	0.0489	0.0489	0.0489	0.0489	0.0489	0.0489	<b>0.0488</b>	0.0488	0.0488	0.9
	0.0512	0.0512	0.0512	0.0512	0.0512	0.0512	0.0512	0.0512	0.0512	0.0512	0.5
11	0.0508	0.0508	0.0508	0.0509	0.0509	0.0509	0.0509	0.0509	0.0509	0.0509	0.7
	0.0488	0.0488	0.0488	0.0488	0.0488	0.0488	0.0488	0.0488	0.0488	0.0488	0.9
12	0.0510	0.0510	0.0510	0.0510	0.0510	0.0510	0.0510	0.0511	0.0510	0.0510	0.5
	0.0488	0.0488	0.0488	0.0488	0.0488	0.0488	0.0488	0.0488	0.0488	0.0488	0.7
13	0.0512	0.0512	0.0512	0.0512	0.0512	0.0513	0.0513	0.0512	0.0512	0.0513	0.9
	0.0512	0.0512	0.0512	0.0512	0.0512	0.0512	0.0512	0.0512	0.0512	0.0513	0.5

Dari tabel 3 terlihat bahwa learning rate 0.05 dengan momentum 0.5 nilai MSE terkecil diperoleh pada epoch ke-12 sebesar 0.0491, sedangkan momentum 0.7 pada epoch ke-88 sebesar 0.0488 dan momentum 0.9 pada epoch ke-5 sebesar 0.0495. Gambar 11 merupakan visualisasi klasifikasi data yang dihasilkan



Gambar 11. Visualisasi klasifikasi data Fungsi Aktivasi Sigmoid Biner Learning Rate 0.05, Momentum 0.5, 0.7, 0.9

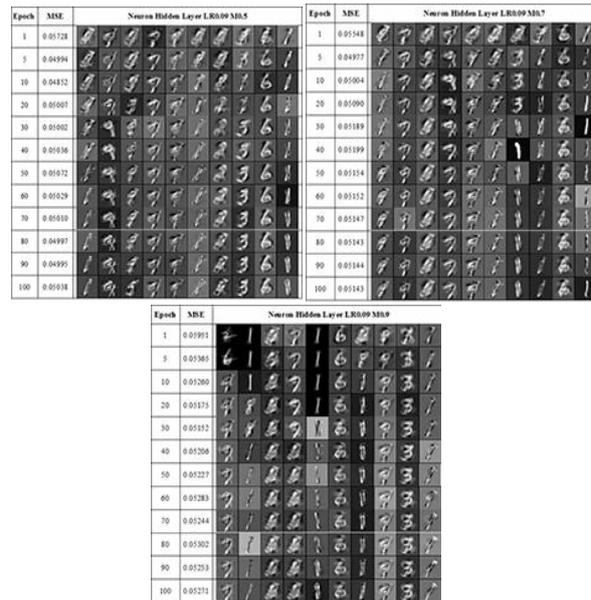
### c. Pembelajaran Tahap III dengan Fungsi Aktivasi Sigmoid Biner

Pada pembelajaran tahap ketiga ini fungsi aktivasi yang digunakan adalah sigmoid biner dengan nilai parameter *learning rate* yang dipilih adalah 0.09 dengan nilai *momentum* bervariasi 0.5, 0.7 dan 0.9, jumlah *epoch* maksimum 100 *epoch*. Hasil lengkap nilai MSE yang diperoleh dapat dilihat pada tabel 4.

Tabel 4. Nilai MSE Fungsi Aktivasi Sigmoid Biner *Learning Rate* 0.09, *Momentum* 0.5, 0.7, 0.9

Ep.	0	1	2	3	4	5	6	7	8	9	Mom.
0	0.0572	0.0501	0.0497	0.0495	0.0499	0.0497	0.0495	<b>0.0485</b>	0.0485	0.0485	<b>0.5</b>
	0.0534	0.0497	<b>0.0492</b>	0.0495	0.0497	0.0498	0.0499	0.0500	0.0500	0.0500	<b>0.7</b>
	0.0395	0.0534	0.0543	0.0540	0.0536	0.0533	0.0525	0.0527	0.0526	0.0526	<b>0.9</b>
1	0.0491	0.0494	0.0495	0.0496	0.0496	0.0496	0.0497	0.0498	0.0498	0.0500	<b>0.5</b>
	0.0501	0.0501	0.0501	0.0502	0.0503	0.0503	0.0505	0.0506	0.0507	0.0509	<b>0.7</b>
	0.0525	0.0525	0.0519	0.0513	0.0514	0.0516	0.0517	0.0517	0.0517	0.0517	<b>0.9</b>
2	0.0501	0.0501	0.0501	0.0502	0.0501	0.0501	0.0500	0.0500	0.0499	0.0500	<b>0.5</b>
	0.0510	0.0511	0.0513	0.0515	0.0516	0.0516	0.0517	0.0518	0.0518	0.0518	<b>0.7</b>
	0.0517	0.0516	0.0516	0.0517	0.0516	0.0515	0.0514	0.0515	0.0515	0.0515	<b>0.9</b>
3	0.0501	0.0500	0.0500	0.0500	0.0500	0.0501	0.0502	0.0502	0.0503	0.0503	<b>0.5</b>
	0.0519	0.0519	0.0519	0.0519	0.0519	0.0519	0.0519	0.0519	0.0519	0.0519	<b>0.7</b>
	0.0515	0.0515	0.0516	0.0516	0.0515	0.0515	<b>0.0512</b>	0.0513	0.0516	0.0520	<b>0.9</b>
4	0.0504	0.0504	0.0505	0.0506	0.0506	0.0507	0.0507	0.0507	0.0507	0.0507	<b>0.5</b>
	0.0519	0.0519	0.0518	0.0517	0.0516	0.0516	0.0515	0.0515	0.0515	0.0515	<b>0.7</b>
	0.0524	0.0524	0.0525	0.0526	0.0526	0.0525	0.0522	0.0518	0.0519	0.0522	<b>0.9</b>
5	0.0507	0.0506	0.0506	0.0505	0.0505	0.0504	0.0504	0.0503	0.0503	0.0502	<b>0.5</b>
	0.0515	0.0515	0.0515	0.0515	0.0515	0.0515	0.0515	0.0515	0.0515	0.0515	<b>0.7</b>
	0.0523	0.0524	0.0525	0.0526	0.0526	0.0527	0.0527	0.0527	0.0528	0.0528	<b>0.9</b>
6	0.0502	0.0502	0.0502	0.0501	0.0501	0.0501	0.0501	0.0501	0.0500	0.0501	<b>0.5</b>
	0.0515	0.0515	0.0515	0.0515	0.0515	0.0515	0.0514	0.0515	0.0514	0.0514	<b>0.7</b>
	0.0528	0.0527	0.0525	0.0524	0.0522	0.0521	0.0521	0.0522	0.0523	0.0524	<b>0.9</b>
7	0.0500	0.0500	0.0500	0.0500	0.0500	0.0499	0.0500	0.0499	0.0499	0.0500	<b>0.5</b>
	0.0514	0.0514	0.0515	0.0515	0.0514	0.0514	0.0515	0.0514	0.0514	0.0514	<b>0.7</b>
	0.0525	0.0527	0.0527	0.0528	0.0529	0.0530	0.0530	0.0530	0.0530	0.0530	<b>0.9</b>
8	0.0499	0.0499	0.0499	0.0499	0.0499	0.0499	0.0499	0.0499	0.0499	0.0499	<b>0.5</b>
	0.0514	0.0514	0.0514	0.0514	0.0514	0.0514	0.0514	0.0514	0.0514	0.0514	<b>0.7</b>
	0.0530	0.0528	0.0525	0.0522	0.0521	0.0522	0.0522	0.0524	0.0524	0.0525	<b>0.9</b>
9	0.0499	0.0499	0.0500	0.0500	0.0501	0.0501	0.0502	0.0502	0.0503	0.0503	<b>0.5</b>
	0.0514	0.0514	0.0514	0.0514	0.0514	0.0514	0.0514	0.0514	0.0514	0.0514	<b>0.7</b>
	0.0527	0.0528	0.0529	0.0529	0.0530	0.0531	0.0531	0.0530	0.0530	0.0527	<b>0.9</b>

Dari tabel 4 terlihat bahwa *learning rate* 0.09 dengan *momentum* 0.5 nilai MSE terkecil diperoleh pada *epoch* ke-8 sebesar 0.0485, sedangkan *momentum* 0.7 pada *epoch* ke-3 sebesar 0.0492 dan *momentum* 0.9 pada *epoch* ke-47 sebesar 0.0512. Gambar 12 berikut merupakan visualisasi klasifikasi data yang dihasilkan



Gambar 12. Visualisasi klasifikasi data Fungsi Aktivasi Sigmoid Biner *Learning Rate* 0.09, *Momentum* 0.5, 0.7, 0.9

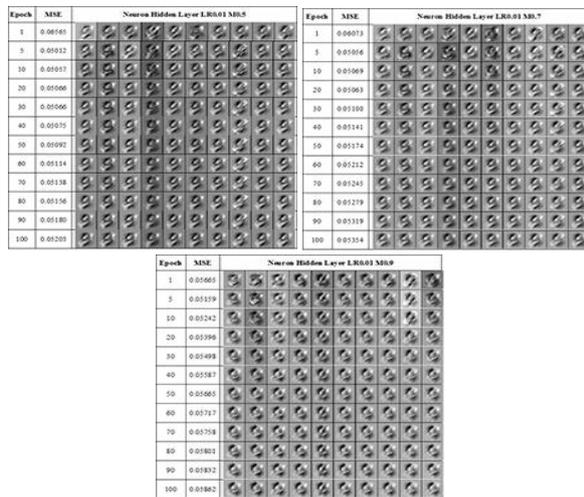
### d. Pembelajaran Tahap IV dengan Fungsi Aktivasi Tanh

Pada pembelajaran tahap keempat ini fungsi aktivasi yang digunakan adalah *tanh* dengan nilai parameter *learning rate* yang dipilih adalah 0.01 dengan nilai *momentum* bervariasi 0.5, 0.7 dan 0.9, jumlah *epoch* maksimum 100 *epoch*. Hasil lengkap nilai MSE yang diperoleh dapat dilihat pada tabel 5.

Tabel 5. Nilai MSE Fungsi Aktivasi Tanh, *Learning Rate* 0.01, *Momentum* 0.5, 0.7, 0.9

Ep.	0	1	2	3	4	5	6	7	8	9	Mom.
0	0.0656	0.0543	0.0521	0.0505	<b>0.0501</b>	0.0502	0.0503	0.0504	0.0505	0.0505	<b>0.5</b>
	0.0607	0.0507	<b>0.0499</b>	0.0503	0.0505	0.0506	0.0507	0.0507	0.0507	0.0506	<b>0.7</b>
	0.0506	0.0517	<b>0.0514</b>	0.0515	0.0515	0.0517	0.0519	0.0521	0.0522	0.0524	<b>0.9</b>
1	0.0506	0.0506	0.0506	0.0506	0.0506	0.0506	0.0506	0.0506	0.0506	0.0506	<b>0.5</b>
	0.0526	0.0527	0.0529	0.0531	0.0533	0.0534	0.0536	0.0537	0.0538	0.0539	<b>0.7</b>
	0.0506	0.0506	0.0506	0.0506	0.0506	0.0506	0.0506	0.0506	0.0506	0.0506	<b>0.9</b>
2	0.0506	0.0506	0.0507	0.0507	0.0508	0.0509	0.0509	0.0509	0.0509	0.0510	<b>0.5</b>
	0.0540	0.0541	0.0542	0.0543	0.0545	0.0545	0.0546	0.0547	0.0549	0.0549	<b>0.7</b>
	0.0506	0.0506	0.0506	0.0506	0.0506	0.0506	0.0506	0.0507	0.0507	0.0507	<b>0.9</b>
3	0.0510	0.0510	0.0511	0.0511	0.0511	0.0512	0.0512	0.0513	0.0513	0.0514	<b>0.5</b>
	0.0551	0.0552	0.0552	0.0554	0.0555	0.0555	0.0556	0.0557	0.0558	0.0558	<b>0.7</b>
	0.0507	0.0507	0.0507	0.0508	0.0508	0.0508	0.0508	0.0508	0.0509	0.0509	<b>0.9</b>
4	0.0514	0.0514	0.0514	0.0515	0.0515	0.0515	0.0516	0.0516	0.0517	0.0517	<b>0.5</b>
	0.0580	0.0580	0.0581	0.0582	0.0582	0.0583	0.0584	0.0584	0.0585	0.0586	<b>0.7</b>
	0.0509	0.0509	0.0510	0.0509	0.0510	0.0510	0.0510	0.0510	0.0510	0.0511	<b>0.9</b>
5	0.0517	0.0518	0.0518	0.0518	0.0519	0.0519	0.0520	0.0520	0.0520	0.0521	<b>0.5</b>
	0.0566	0.0567	0.0567	0.0568	0.0569	0.0569	0.0570	0.0570	0.0570	0.0571	<b>0.7</b>
	0.0511	0.0511	0.0511	0.0512	0.0512	0.0512	0.0512	0.0513	0.0513	0.0513	<b>0.9</b>
6	0.0521	0.0521	0.0522	0.0522	0.0523	0.0523	0.0523	0.0523	0.0524	0.0524	<b>0.5</b>
	0.0571	0.0572	0.0572	0.0573	0.0573	0.0574	0.0575	0.0575	0.0575	0.0575	<b>0.7</b>
	0.0513	0.0514	0.0514	0.0514	0.0514	0.0514	0.0515	0.0515	0.0515	0.0515	<b>0.9</b>
7	0.0524	0.0525	0.0525	0.0526	0.0526	0.0526	0.0527	0.0527	0.0527	0.0527	<b>0.5</b>
	0.0576	0.0576	0.0577	0.0577	0.0578	0.0578	0.0579	0.0579	0.0580	0.0580	<b>0.7</b>
	0.0515	0.0516	0.0516	0.0516	0.0516	0.0517	0.0517	0.0517	0.0517	0.0518	<b>0.9</b>
8	0.0528	0.0528	0.0529	0.0529	0.0529	0.0530	0.0530	0.0531	0.0531	0.0531	<b>0.5</b>
	0.0580	0.0580	0.0581	0.0581	0.0581	0.0582	0.0582	0.0582	0.0583	0.0583	<b>0.7</b>
	0.0518	0.0518	0.0518	0.0518	0.0519	0.0519	0.0519	0.0519	0.0519	0.0520	<b>0.9</b>
9	0.0532	0.0532	0.0532	0.0533	0.0533	0.0534	0.0534	0.0534	0.0535	0.0535	<b>0.5</b>
	0.0583	0.0584	0.0584	0.0584	0.0584	0.0585	0.0585	0.0585	0.0585	0.0586	<b>0.7</b>
	0.0518	0.0518	0.0518	0.0518	0.0518	0.0518	0.0518	0.0518	0.0518	0.0518	<b>0.9</b>

Dari tabel 5 terlihat bahwa fungsi aktivasi *tanh* dengan *learning rate* 0.01 dan *momentum* 0.5 *epoch* ke-5 memiliki nilai MSE terkecil dengan nilai 0.0501, sedangkan *momentum* 0.7 pada *epoch* ke-3 sebesar 0.0499 dan *momentum* 0.9 pada *epoch* ke-3 sebesar 0.0514. Gambar 13 berikut merupakan visualisasi klasifikasi data yang dihasilkan.



Gambar 13. Visualisasi klasifikasi data Fungsi Aktivasi *Tanh*, *Learning Rate* 0.01, *Momentum* 0.5, 0.7, 0.9

**e. Pembelajaran Tahap V dengan Fungsi Aktivasi *Tanh***

Pada pembelajaran tahap kelima ini fungsi aktivasi yang digunakan adalah *tanh* dengan nilai parameter *learning rate* yang dipilih adalah 0.05 dengan nilai *momentum* bervariasi 0.5, 0.7 dan 0.9, jumlah *epoch* maksimum 100 *epoch*. Hasil lengkap nilai MSE yang diperoleh dapat dilihat pada tabel 6.

Tabel 6. Nilai MSE Fungsi Aktivasi *Tanh*, *Learning Rate* 0.05, *Momentum* 0.5, 0.7, 0.9

Ep.	0	1	2	3	4	5	6	7	8	9	Mom.
0	0.0538	0.0505	0.0506	0.0505	0.0504	0.0504	0.0504	0.0505	0.0506	0.0507	0.5
	0.0538	0.0511	0.0512	0.0514	0.0516	0.0518	0.0521	0.0525	0.0526	0.0528	0.7
1	0.0612	0.0590	0.0596	0.0601	0.0604	0.0608	0.0612	0.0614	0.0618	0.0620	0.9
	0.0598	0.0510	0.0510	0.0512	0.0513	0.0514	0.0515	0.0516	0.0518	0.0519	0.5
2	0.0530	0.0532	0.0535	0.0537	0.0539	0.0541	0.0543	0.0545	0.0547	0.0550	0.7
	0.0622	0.0624	0.0626	0.0627	0.0629	0.0629	0.0631	0.0633	0.0634	0.0634	0.9
3	0.0520	0.0521	0.0522	0.0524	0.0524	0.0525	0.0526	0.0529	0.0529	0.0529	0.5
	0.0551	0.0552	0.0554	0.0555	0.0556	0.0557	0.0558	0.0559	0.0560	0.0561	0.7
4	0.0635	0.0637	0.0638	0.0638	0.0640	0.0640	0.0642	0.0643	0.0644	0.0645	0.9
	0.0531	0.0532	0.0533	0.0534	0.0535	0.0536	0.0537	0.0538	0.0539	0.0540	0.5
5	0.0562	0.0563	0.0564	0.0565	0.0565	0.0566	0.0567	0.0568	0.0569	0.0569	0.7
	0.0646	0.0646	0.0647	0.0648	0.0649	0.0649	0.0650	0.0651	0.0651	0.0653	0.9
6	0.0541	0.0542	0.0543	0.0544	0.0545	0.0546	0.0546	0.0547	0.0548	0.0549	0.5
	0.0570	0.0571	0.0571	0.0572	0.0573	0.0574	0.0575	0.0575	0.0576	0.0576	0.7
7	0.0652	0.0653	0.0654	0.0655	0.0655	0.0655	0.0655	0.0656	0.0657	0.0657	0.9
	0.0517	0.0517	0.0518	0.0518	0.0519	0.0519	0.0520	0.0521	0.0521	0.0522	0.5
8	0.0638	0.0638	0.0639	0.0639	0.0639	0.0640	0.0640	0.0641	0.0642	0.0642	0.7
	0.0535	0.0535	0.0536	0.0536	0.0537	0.0537	0.0538	0.0538	0.0539	0.0539	0.9
9	0.0582	0.0583	0.0583	0.0583	0.0584	0.0584	0.0585	0.0585	0.0585	0.0586	0.5
	0.0642	0.0643	0.0643	0.0644	0.0644	0.0644	0.0644	0.0645	0.0645	0.0645	0.7
10	0.0580	0.0580	0.0581	0.0581	0.0582	0.0582	0.0583	0.0583	0.0583	0.0584	0.9
	0.0586	0.0587	0.0587	0.0588	0.0588	0.0589	0.0589	0.0589	0.0590	0.0590	0.5
11	0.0665	0.0666	0.0666	0.0666	0.0666	0.0666	0.0667	0.0667	0.0667	0.0667	0.7
	0.0564	0.0565	0.0565	0.0565	0.0566	0.0566	0.0566	0.0567	0.0567	0.0567	0.9
12	0.0590	0.0590	0.0591	0.0591	0.0592	0.0592	0.0593	0.0593	0.0594	0.0594	0.5
	0.0667	0.0667	0.0668	0.0668	0.0668	0.0668	0.0668	0.0668	0.0669	0.0669	0.7
13	0.0568	0.0568	0.0568	0.0569	0.0569	0.0569	0.0570	0.0570	0.0571	0.0571	0.9
	0.0595	0.0595	0.0595	0.0595	0.0596	0.0596	0.0597	0.0597	0.0597	0.0597	0.5
14	0.0669	0.0670	0.0670	0.0670	0.0670	0.0670	0.0671	0.0671	0.0670	0.0671	0.7
	0.0598	0.0598	0.0598	0.0598	0.0598	0.0598	0.0598	0.0598	0.0598	0.0598	0.9

Dari tabel 6 terlihat bahwa fungsi aktivasi *tanh* dengan *learning rate* 0.05 dan *momentum* 0.5 *epoch* ke-5 memiliki nilai MSE terkecil dengan nilai 0.0504, sedangkan *momentum* 0.7 pada *epoch* ke-2 sebesar 0.0511 dan *momentum* 0.9 pada *epoch* ke-2 sebesar 0.0590. Visualisasi klasifikasi data yang dihasilkan cenderung sama dengan gambar 13.

**f. Pembelajaran Tahap VI dengan Fungsi Aktivasi *Tanh***

Pada pembelajaran tahap keenam ini fungsi aktivasi yang digunakan adalah *tanh* dengan nilai parameter *learning rate* yang dipilih adalah 0.09 dengan nilai *momentum* bervariasi 0.5, 0.7 dan 0.9, jumlah *epoch* maksimum 100 *epoch*. Hasil lengkap nilai MSE yang diperoleh dapat dilihat pada tabel 7.

Tabel 7 Nilai MSE Fungsi Aktivasi *Tanh* *LearningRate* 0.09, *Momentum* 0.5, 0.7, 0.9

Ep.	0	1	2	3	4	5	6	7	8	9	Mem.
0	0.0532	0.0508	0.0508	0.0508	0.0510	0.0512	0.0514	0.0515	0.0517	0.0519	0.5
0	0.0536	0.0518	0.0521	0.0525	0.0528	0.0533	0.0538	0.0541	0.0544	0.0548	0.7
0	0.0639	0.0665	0.0669	0.0674	0.0678	0.0685	0.0685	0.0687	0.0688	0.0691	0.9
1	0.0530	0.0523	0.0525	0.0528	0.0530	0.0533	0.0536	0.0539	0.0542	0.0545	0.5
1	0.0531	0.0534	0.0537	0.0539	0.0541	0.0543	0.0545	0.0547	0.0548	0.0550	0.7
1	0.0692	0.0694	0.0695	0.0697	0.0699	0.0698	0.0700	0.0701	0.0703	0.0703	0.9
2	0.0538	0.0540	0.0541	0.0543	0.0544	0.0545	0.0546	0.0547	0.0548	0.0549	0.5
2	0.0571	0.0573	0.0574	0.0576	0.0577	0.0578	0.0579	0.0580	0.0581	0.0582	0.7
2	0.0705	0.0705	0.0705	0.0698	0.0684	0.0687	0.0690	0.0692	0.0693	0.0693	0.9
3	0.0530	0.0531	0.0532	0.0533	0.0534	0.0535	0.0536	0.0537	0.0538	0.0539	0.5
3	0.0583	0.0584	0.0585	0.0585	0.0586	0.0587	0.0588	0.0588	0.0589	0.0589	0.7
3	0.0696	0.0696	0.0698	0.0697	0.0698	0.0700	0.0701	0.0702	0.0702	0.0702	0.9
4	0.0590	0.0591	0.0591	0.0592	0.0593	0.0594	0.0594	0.0595	0.0596	0.0596	0.5
4	0.0590	0.0591	0.0592	0.0593	0.0593	0.0594	0.0594	0.0595	0.0595	0.0596	0.7
4	0.0702	0.0703	0.0704	0.0704	0.0705	0.0706	0.0706	0.0706	0.0706	0.0706	0.9
5	0.0567	0.0568	0.0568	0.0569	0.0570	0.0570	0.0570	0.0571	0.0572	0.0572	0.5
5	0.0596	0.0597	0.0598	0.0598	0.0599	0.0600	0.0600	0.0601	0.0601	0.0602	0.7
5	0.0708	0.0707	0.0707	0.0708	0.0708	0.0709	0.0710	0.0710	0.0708	0.0709	0.9
6	0.0573	0.0573	0.0574	0.0574	0.0575	0.0575	0.0576	0.0577	0.0577	0.0578	0.5
6	0.0602	0.0603	0.0603	0.0603	0.0604	0.0604	0.0605	0.0605	0.0605	0.0606	0.7
6	0.0709	0.0711	0.0711	0.0711	0.0711	0.0711	0.0711	0.0711	0.0713	0.0713	0.9
7	0.0578	0.0579	0.0579	0.0580	0.0580	0.0581	0.0581	0.0582	0.0582	0.0583	0.5
7	0.0606	0.0606	0.0607	0.0607	0.0607	0.0607	0.0608	0.0609	0.0609	0.0609	0.7
7	0.0713	0.0713	0.0713	0.0712	0.0713	0.0713	0.0713	0.0714	0.0714	0.0714	0.9
8	0.0583	0.0583	0.0584	0.0584	0.0584	0.0585	0.0585	0.0586	0.0586	0.0587	0.5
8	0.0610	0.0610	0.0610	0.0610	0.0610	0.0611	0.0611	0.0611	0.0611	0.0612	0.7
8	0.0714	0.0714	0.0714	0.0714	0.0715	0.0714	0.0715	0.0715	0.0715	0.0715	0.9
9	0.0587	0.0587	0.0587	0.0587	0.0588	0.0588	0.0588	0.0589	0.0589	0.0590	0.5
9	0.0612	0.0612	0.0612	0.0612	0.0613	0.0613	0.0614	0.0614	0.0614	0.0614	0.7
9	0.0715	0.0715	0.0715	0.0716	0.0715	0.0715	0.0716	0.0716	0.0716	0.0716	0.9

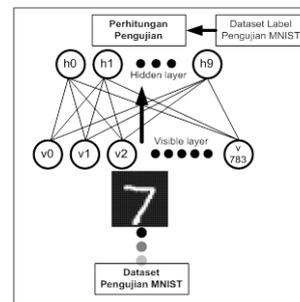
Dari tabel 7 terlihat bahwa fungsi aktivasi *tanh* dengan *learning rate* 0.09 dan *momentum* 0.5 *epoch* ke-2 memiliki nilai MSE terkecil dengan nilai 0.0508, sedangkan momentum 0.7 pada epoch ke-2 sebesar 0.0518 dan momentum 0.9 pada epoch ke-1 sebesar 0.0659. Visualisasi klasifikasi data yang dihasilkan cenderung sama dengan gambar 13.

Berdasarkan tahapan pembelajaran tersebut dapat ditarik kesimpulan bahwa pada pembelajaran yang menggunakan fungsi aktivasi sigmoid dengan pilihan parameter *learning rate* dan momentum yang tepat dapat menurunkan nilai MSE menjadi lebih kecil sehingga akan mempengaruhi hasil klasifikasi angka tulisan tangan. Sedangkan penggunaan fungsi aktivasi *tanh* pada jaringan RBM terlihat tidak mampu menurunkan nilai MSE di bawah fungsi aktivasi sigmoid. Nilai MSE yang dihasilkan cenderung menaik seiring meningkatnya jumlah epoch. Nilai MSE yang cenderung menurun hingga akhir epoch akan menghasilkan visualisasi klasifikasi tulisan tangan yang jelas. Hasil

visualisasi klasifikasi angka tulisan tangan ini dapat dilihat lebih jelas pada bagian pembahasan.

### 2.2.4 Pengujian

Proses pengujian dilakukan dengan memasukkan nilai-nilai bobot akhir yang akan diuji ke dalam jaringan RBM, dilanjutkan dengan memberi masukan dataset pengujian sebanyak 10000 data angka tulisan tangan. Pada proses pengujian, jaringan RBM hanya memproses dari visible layer ke hidden layer (phase positif) saja. Hasil perhitungan pada hidden layer selanjutnya dihitung dan diklasifikasikan berdasarkan data label MNIST untuk memperoleh akurasi pengenalan.



Gambar 14. Arsitektur Pengujian Dataset MNIST

Pengujian dilakukan dengan cara mengisi langsung semua variabel-variabel bobot pada jaringan RBM dengan nilai bobot terakhir dari proses pembelajaran yang telah disimpan sebelumnya. Jumlah file yang berisi nilai-nilai bobot terdiri atas 100 file, sesuai dengan jumlah maksimum *epoch* saat pembelajaran. Pengambilan nilai-nilai bobot akhir dari file ini sangat

menghemat waktu pembelajaran sehingga tidak diperlukan pembelajaran ulang jaringan.

Pengujian dimulai dengan membaca file-file bobot setiap *epoch* untuk *learning rate* 0.01 dengan *momentum* 0.5. Hasil perhitungan pengenalan angka tulisan tangan setiap *epoch* disimpan untuk keperluan analisa selanjutnya. Pengujian yang sama dilanjutkan untuk untuk *learning rate* 0.01, *momentum* 0.7 dan 0.9 dan *learning rate* 0.05 dengan *momentum* 0.5 dan 0.7. Hasil lengkap pengujian untuk *learning rate* 0.01 dan 0.05 masing-masing dapat dilihat pada tabel 8 dan 9. Isi tabel berupa jumlah data angka tulisan tangan yang cocok dibandingkan dengan label dataset pada data pengujian.

Tabel 8. Hasil Pengujian *Learning Rate* 0.01, *Momentum* 0.5, 0.7, 0.9

Ep	0	1	2	3	4	5	6	7	8	9	Mom
0	1628	1818	1188	1857	2838	3595	4147	4462	4583	4614	0.5
	1615	2424	4147	4886	4975	5941	5284	5654	6070	6439	0.7
	2527	4732	5187	5713	6289	6876	7395	7715	7952	8110	0.9
1	4615	4658	4719	4787	4866	4970	5092	5234	5393	5573	0.5
	6748	7653	7277	7456	7838	7770	7881	8001	8070	8159	0.7
	8243	8356	8443	8514	8594	8640	8679	8709	8744	8764	0.9
2	5775	5965	6109	6281	6414	6580	6709	6845	6979	7069	0.5
	8219	8267	8331	8372	8410	8447	8484	8522	8552	8588	0.7
	8889	8827	8850	8865	8878	8890	8904	8915	8929	8940	0.9
3	7153	7248	7339	7418	7492	7568	7625	7688	7749	7815	0.5
	8618	8643	8669	8689	8710	8727	8750	8769	8777	8794	0.7
	8951	8970	8973	8980	8981	8984	8988	8995	8999	9005	0.9
4	7859	7980	7941	7969	8022	8060	8085	8108	8135	8162	0.5
	8889	8818	8823	8829	8840	8850	8854	8860	8865	8875	0.7
	9012	9014	9017	9028	9033	9044	9046	9046	9052	9056	0.9
5	8181	8209	8228	8254	8270	8290	8299	8314	8338	8350	0.5
	8881	8887	8895	8899	8905	8911	8913	8915	8919	8924	0.7
	9062	9061	9065	9069	9076	9081	9083	9088	9094	9103	0.9
6	8352	8381	8396	8411	8428	8442	8459	8477	8488	8500	0.5
	8828	8941	8946	8949	8953	8954	8961	8967	8975	8974	0.7
	9186	9108	9112	9114	9115	9120	9124	9126	9129	9134	0.9
7	8511	8521	8529	8548	8556	8570	8580	8589	8595	8600	0.5
	8974	8980	8983	8985	8989	8992	8995	8995	8998	9000	0.7
	9137	9139	9146	9150	9151	9153	9153	9155	9156	9157	0.9
8	8609	8620	8633	8638	8648	8654	8662	8671	8679	8678	0.5
	9001	9007	9012	9016	9018	9019	9021	9021	9024	9028	0.7
	9165	9167	9169	9171	9173	9173	9175	9175	9176	9175	0.9
9	8681	8693	8701	8717	8721	8723	8731	8738	8746	8749	0.5
	9027	9032	9033	9034	9036	9039	9044	9048	9051	9056	0.7
	9176	9178	9180	9181	9184	9185	9188	9188	9192	9191	0.9

Tabel 8 memperlihatkan hasil pengujian pengenalan 10000 data angka untuk masing-masing parameter, *learning rate* 0.01 dengan *momentum* 0.5, 0.7 dan 0.9, dimulai dari *epoch* 1 sampai *epoch* 100.

Tabel 9. Hasil Pengujian *Learning Rate* 0.05, *Momentum* 0.5, 0.7

Ep	0	1	2	3	4	5	6	7	8	9	Mom
0	3857	5684	4934	5643	6488	7127	7513	7859	8057	8282	0.5
	5227	6583	7640	8145	8414	8594	8719	8759	8841	8922	0.7
	8380	8386	8463	8537	8591	8625	8666	8702	8731	8748	0.9
1	8905	8932	8960	8977	8993	9012	9025	9035	9044	9056	0.7
	8767	8781	8797	8815	8825	8844	8860	8872	8883	8909	0.5
	9066	9073	9079	9093	9094	9099	9105	9119	9126	9132	0.7
2	8921	8928	8931	8935	8946	8953	8956	8963	8963	8975	0.5
	9139	9142	9146	9152	9155	9158	9165	9167	9173	9176	0.7
	8981	8989	8990	8990	8993	8998	9005	9006	9004	9014	0.5
3	9177	9178	9180	9188	9189	9196	9197	9201	9205	9209	0.7
	9020	9024	9028	9036	9040	9044	9047	9052	9055	9061	0.5
	9011	9010	9013	9019	9020	9023	9022	9025	9032	9035	0.7
4	9066	9065	9066	9067	9068	9070	9072	9076	9076	9076	0.5
	9235	9237	9243	9247	9250	9253	9251	9250	9254	9262	0.7
	9081	9085	9085	9088	9088	9091	9095	9096	9097	9100	0.5
5	9154	9209	9271	9274	9270	9281	9282	9285	9285	9289	0.7
	9186	9186	9110	9113	9114	9114	9117	9117	9119	9123	0.5
	9291	9296	9297	9300	9304	9307	9310	9311	9313	9315	0.7
6	9124	9125	9126	9125	9124	9128	9130	9130	9131	9131	0.5
	9322	9324	9326	9329	9334	9338	9339	9340	9342	9342	0.7

Tabel 9 memperlihatkan hasil pengujian pengenalan 10000 data angka untuk masing-masing parameter, *learning rate* 0.05 dengan *momentum* 0.5, dan 0.7 dimulai dari *epoch* 1 sampai *epoch* 100.

Tabel 8 dan 9 memperlihatkan secara rinci persentase yang dihasilkan pada proses pengujian. Terlihat parameter *learning rate* 0.01 dengan *momentum* 0.5 dapat mengenali dataset pengujian tulisan tangan maksimal 8749 atau sama artinya dengan 87.49%. Dilanjutkan dengan *learning rate* 0.01 dengan *momentum* 0.7 dapat mengenal dataset pengujian maksimum sebesar 9056 atau sama artinya dengan 90.56%. Berikutnya *learning rate* 0.01 dengan *momentum* 0.9 dan *learning rate* 0.05 dengan *momentum* 0.5, memiliki kinerja yang hampir sama, dapat mengenali dataset masing-masing 9192 sama artinya dengan 91.92% dan 9131 sama artinya dengan 91.31%. Terakhir, parameter *learning rate* 0.05 dengan *momentum* 0.7 memiliki kinerja yang paling tinggi, mampu mengenali

dataset pengujian sebesar 9342 atau sama artinya dengan 93.42%.

## SIMPULAN

Kinerja RBM sangat dipengaruhi oleh pemilihan nilai dari parameter learning rate, momentum dan fungsi aktivasi yang tepat, khususnya dalam menentukan nilai *mean square error* (MSE) dan persentase akurasi pengenalan terhadap dataset. Pembelajaran yang menggunakan fungsi aktivasi sigmoid, tingkat keberhasilan klasifikasi ditentukan oleh nilai MSE yang kecil. Sedangkan penggunaan fungsi aktivasi tanh pada jaringan RBM ternyata tidak dapat menurunkan nilai MSE di bawah nilai yang dihasilkan oleh fungsi aktivasi sigmoid, nilai MSE menaik seiring meningkatnya jumlah epoch.

Dari hasil pembahasan terlihat bahwa pemakaian fungsi aktivasi sigmoid dengan nilai learning rate tinggi dan momentum yang rendah memiliki tingkat persentase pengenalan dataset angka tulisan tangan yang tinggi.

## UCAPAN TERIMAKASIH

Penulis mengucapkan banyak terima kasih kepada Universitas Medan Area (UMA) yang telah memberikan dukungan dan memfasilitasi kegiatan penelitian ini.

## DAFTAR PUSTAKA

Ahmed. A, Yu. K, Xu. W, Gong. Y, dan Xing. P. E (2008). Training hierarchical feed-forward visual recognition models using transfer

learning from pseudo tasks, *Proceedings of the 10th European Conference on Computer Vision (ECCV'08)*: pp. 69–82.

- Bengio. Y, Lamblin. P, Popovici. D, dan Larochelle. H (2007). Greedy layer-wise training of deep networks, *Advances in Neural Information Processing Systems 19 (NIPS'06)* : pp. 153–160
- Collobert. R, dan Weston. J (2008). A unified architecture for natural language processing: Deep neural networks with multitask learning, *Proceedings of the Twenty-fifth International Conference on Machine Learning (ICML'08)*, pp. 160–167.
- Hadsell. R, Erkan. A, Sermanet. P, Scoffier. M, Muller. U, dan LeCun. Y (2008). Deep belief net learning in a long-range vision system for autonomous offroad driving, *Proc. Intelligent Robots and Systems (IROS'08)*, pp. 628–633.
- Hinton, Geoffrey. 2010. *A Practical Guide to Training Restricted Boltzmann Machines*. University of Toronto.
- Hinton. E. G, Osindero. S, dan Teh. W. Y (2006). A fast learning algorithm for deep belief nets. *Neural Computation*, 18(7):1527–1554.
- Hinton. E. G (2002). Training products of experts by minimizing contrastive divergence, *Neural Computation*, vol. 14, pp. 1771–1800.
- Kai, Ding., Zhibin, Liu., Lianwen, Jin. & Xinghua, Zhu. 2007. A Comparative study of GABOR feature and gradient feature for handwritten Chinese character recognition, *International Conference on Wavelet Analysis and Pattern Recognition* : pp. 1182–1186
- Larochelle. H, Erhan. D, Courville. A, Bergstra. J, dan Bengio. Y (2007). An empirical evaluation of deep architectures on problems with many factors of variation, *Proceedings of the Twenty-fourth International Conference on Machine Learning (ICML'07)* : pp. 473–480.
- Le Cun, Yann. & Corinna Cortes. 2010. *The MNIST Database of Handwritten Digits*. Web. <<http://yann.lecun.com/exdb/mnist/>>.
- Lee. H, Grosse. R, Ranganath. R, dan Ng. Y. A (2009). Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations, *Proceedings of the Twenty-sixth International Conference on Machine Learning (ICML'09)*, Montreal (Qc), Canada: ACM.
- Lotfi, Abdelhadi. & Benyettou, Abdelkader (2011). Using probabilistic neural network for handwritten digit recognition. *Journal of Artificial Intelligence*. ISSN 1994-5450.
- Mnih. A, dan Hinton. E. G (2009). A scalable hierarchical distributed language model, *Advances in Neural Information Processing Systems 21 (NIPS'08)* : pp. 1081–1088.

- Osindero. S, dan Hinton. E. G (2008). Modeling image patches with a directed hierarchy of Markov random field, *Advances in Neural Information Processing Systems 20 (NIPS'07)*: pp. 1121–1128.
- Ranzato. M, Boureau. -L. Y, dan LeCun. Y (2008). Sparse feature learning for deep belief networks, in *Advances in Neural Information Processing Systems 20 (NIPS'07)*: pp. 1185–1192.
- Ranzato. M, Poultney. C, Chopra. S, dan LeCun. Y (2007). Efficient learning of parse representations with an energy-based model, *Advances in Neural Information Processing Systems 19 (NIPS'06)*: pp. 1137–1144.
- Ranzato. M, dan Szummer. M (2008). Semi-supervised learning of compact document representations with deep networks, *Proceedings of the Twenty-fifth International Conference on Machine Learning (ICML'08)*: vol. 307, pp. 792–799.
- Salakhutdinov. R, dan Hinton. E. G (2008). Using deep belief nets to learn covariance kernels for Gaussian processes, *Advances in Neural Information Processing Systems 20 (NIPS'07)*: pp. 1249–1256.
- Salakhutdinov. R, dan Hinton. E. G (2007). Learning a nonlinear embedding by preserving class neighbourhood structure, *Proceedings of the Eleventh International Conference on Artificial Intelligence and Statistics (AISTATS'07)*: San Juan, Porto Rico: Omnipress.
- Salakhutdinov. R, dan Hinton. E. G (2007). Semantic hashing. *Proceedings of the 2007 Workshop on Information Retrieval and applications of Graphical Models (SIGIR 2007)*, Amsterdam: Elsevier.
- Salakhutdinov. R, Mnih. A, dan Hinton E. G (2007). Restricted Boltzmann machines for collaborative filtering, *Proceedings of the Twenty-fourth International Conference on Machine Learning (ICML'07)*: pp. 791–798, New York.
- Smolensky. P (1986). Information processing in dynamical systems: *Foundations of harmony theory, Parallel Distributed Processing*, vol. 1, pp. 194–28.
- Susilawati, (2017), Algoritma Restricted Boltzmann Machines (RBM) untuk Pengenalan Tulisan Tangan Angka, Prosiding seminar nasional Teknologi Informatika, ISBN :978-602-50006-0-7, pp. 128-136
- Taylor. G, dan Hinton. G (2009). Factored conditional restricted Boltzmann machines for modeling motion style, *Proceedings of the 26th International Conference on Machine Learning (ICML'09)*: pp. 1025–1032.
- Taylor. G, Hinton. E. G, dan Roweis. S (2007). Modeling human motion using binary latent variables, *Advances in Neural Information Processing Systems 19 (NIPS'06)*: pp. 1345–1352.
- Torralba. A, Fergus. R, dan Weiss. Y (2008). Small codes and large databases for recognition, *Proceedings of the Computer Vision and Pattern Recognition Conference (CVPR'08)*: pp. 1–8.
- Weston. J, Ratle. F, dan Collobert. R (2008). Deep learning via semi-supervised embedding, *Proceedings of the Twenty-fifth International Conference on Machine Learning (ICML'08)*: pp. 1168–1175, New York.
- Yee, The. & Hinton, Geoffrey. (2010). Rate-coded Restricted Boltzmann Machines for Face Recognition. University of Toronto. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.135.5929&rep=rep1&type=pdf>>. 12 September 2014.