

Self Organizing Maps (Kohonen) untuk Cluster Bidang Karya Ilmiah (Skripsi) Mahasiswa Berdasarkan Nilai-Nilai Matakuliah Pendukung Machine Learning

Hery Sunandar, Yasir Hasan*

Fakultas Ilmu Komputer Dan Teknologi Informasi, Prodi Teknik Informatika, Universitas Budi Darma, Medan, Indonesia

Email: ¹herysun1975@gmail.com, ^{2,*}yasirhasan.kom@gmail.com

Submitted: 14/08/2022; Accepted: 24/08/2022; Published: 30/08/2022

Abstrak—Waktu penyusunan karya ilmiah dianggap terlalu cepat dan tidak sesuai bagi mahasiswa dikarenakan beberapa hal dalam ruanglingkup penyusunannya. Salah satu alasannya mahasiswa tersebut tidak mampu menyelesaikan dan aspek pendukung bidang karya ilmiah yang dikerjakan, banyak mahasiswa yang topik karya ilmiahnya berada pada *machine learning* atau sejenisnya, namun mahasiswa tersebut tidak mampu menyelesaikan karya ilmiahnya dikarenakan tidak memahami teoritis dari pengetahuan pendukung dari bidang *machine learning*. Sehingga dapat membuat mahasiswa tersebut depresi atau lebih beratnya mengulang karya ilmiahnya semester selanjutnya. Ruanglingkup matakuliah *machine learning* adalah matakuliah statistik dan probabilitas, matrik dan aljabar linear, algoritma dan pemrograman, dan struktur data. penelitian ini dilakukan untuk menanggulangi masalah-masalah yang dihadapi mahasiswa yaitu mengetahui kelompok kesesuaian bidang karya ilmiah mahasiswa yang akan mereka kerjakan. Sehingga dalam penyusunannya mahasiswa dapat bertanggung jawab dalam sidangnya dan lebih berkualitas. Pengujian *clustering* dengan algoritma *self organizing maps* (SOM) lebih stabil dikarenakan untuk input sesuai data yang dimiliki bukan random, hanya pembobotan saja yang dilakukan secara *random* namun berdasarkan nilai batas *uniform low* (mins) dan *high* (maks). Jumlah *cluster* yang diinginkan ada dua yaitu *cluster 0* mampu mengerjakan karya ilmiah berbasis *machine learning* dan *cluster 1* sebaliknya. Proses SOM terhadap 40 data mahasiswa dengan target sebanyak dua *cluster* dan hasilnya *cluster 0* = 14 mahasiwa, *cluster 1* = 26 mahasiswa. Hasilnya didapatkan dengan menaikkan radius = 1, yang mana sebelumnya pencapaian ini tidak berhasil jika radius = 0.

Kata Kunci: Self Organizing Maps; Kohonen Maps; Clustering; Tugas Akhir; Machine Learning

Abstract—The time for the preparation of scientific papers is considered too fast and not suitable for students due to several things in the scope of its preparation. One of the reasons is that the student is unable to complete and the supporting aspects of the field of scientific work being done, many students whose topics of scientific work are in machine learning or the like, but these students are unable to complete their scientific work because they do not understand the theoretical supporting knowledge of the machine learning field. So that it can make the student depressed or even harder to repeat his scientific work the next semester. The scope of machine learning courses are statistics and probability, matrix and linear algebra, algorithms and programming, and data structures. This research was conducted to overcome the problems faced by students, namely knowing the suitability group for the field of student scientific work they will be working on. So that in its preparation students can be responsible in their trial and are of higher quality. The clustering test with the self organizing maps (SOM) algorithm is more stable because the input according to the data owned is not random, only the weighting is done randomly but based on the uniform low (mins) and high (max) limit values. The desired number of clusters is two namely cluster 0 is able to do scientific work based on machine learning and cluster 1 vice versa. The SOM process for 40 student data with a target of two clusters and the results are cluster 0 = 14 students, cluster 1 = 26 students. The result is obtained by increasing the radius = 1, which previously this achievement was not successful if radius = 0.

Keywords: Self Organizing Maps; Kohonen Maps; Clustering; Final Project; Machine Learning

1. PENDAHULUAN

Proyek karya ilmiah (skripsi) kegiatan akhir sebagai penentu kelulusan mahasiswa yang telah melaksanakan pendidikan di kampusnya. Periode pelaksanaan penyusunan karya ilmiah mahasiswa dilakukan selama satu semester (enam bulan) di tahun ajaran genap[1], [2]. Waktu penyusunan karya ilmiah dianggap terlalu cepat dan tidak sesuai bagi mahasiswa dikarenakan beberapa hal dalam ruanglingkup penyusunannya [3]. Salah satu alasannya mahasiswa tersebut tidak memahami cara penyelesaian dan aspek pendukung bidang karya ilmiah yang dikerjakan, seperti banyak mahasiswa yang topik karya ilmiahnya berada pada *machine learning* atau sejenis bidang tersebut, namun mahasiswa tersebut tidak mampu menyelesaikan karya ilmiahnya dikarenakan tidak memahami teoritis dari pengetahuan pendukung dari bidang *machine learning*. Sehingga dapat membuat mahasiswa tersebut depresi atau lebih beratnya mengulang karya ilmiahnya semester selanjutnya. Karya ilmiah yang sesuai ruang lingkup matakuliah *machine learning* adalah matakuliah statistik dan probabilitas, matrik dan aljabar linear, struktur data, algoritma dan pemrograman dan matakuliah pemrograman untuk perancangan dan implementasi aplikasinya [4]. Setiap materi yang disuguhkan memiliki pemecahan masalah yang terdapat matakuliah pendukung tersebut dalam mewujudkan algoritma yang digunakan *machine learning* baik dalam *supervised learning* maupun *unsupervised learning* [5], [6].

Clustering terdapat pada bagian *machine learning* yang dikenal dengan *unsupervised learning* [7][8]. Selain di *machine learning*, *clustering* juga dipelajari di *data mining*, karena *data mining* dan *machine learning* sangat berkaitan erat dalam *artificial intelligence* (AI) [9], [10]. Jadi *clustering* memiliki fungsi sebagai pengelompokkan dari data berdasarkan sifat atau karakteristik masing-masing. Hal ini disebut kelompok data

yang similar atau juga disimilar [11], [12]. Pengurutan data dilakukann berdasarkan karakteristi dengan kemiripan paling dekat dengan data lainnya akan dikelompokkan menjadi satu klaster. Demikian upaya ini berguna untuk mengarahkan hasil temuan *cluster* yang sebelumnya tidak diketahui pada data [13], [14].

Clustering bidang karya ilmiah mahasiswa dilakukan untuk menanggulangi masalah-masalah yang dihadapi mahasiswa. Cara ini dilakukan untuk mengetahui kelompok kesesuaian bidang karya ilmiah mahasiswa yang akan mereka kerjakan. Sehingga dalam penyusunan karya ilmiah mahasiswa tersebut dapat bertanggung jawab dalam sidangnya dan karya ilmiah mereka juga lebih berkualitas. Dalam pengujian ini digunakan *clustering* dengan algoritma *self organizing maps* (SOM) / kohonen maps diketahui SOM lebih stabil dikarenakan untuk input sesuai data yang dimiliki bukan *random*, hanya pembobotan saja yang dilakukan secara *random* namun berdasarkan nilai batas. Penelitian *clustering* dengan SOM pernah dilakukan Mujiati Dwi Kartikasari menerapkan SOM dan Davies-Bouldin Index untuk mengelompokkan wilayah Indonesia berdasarkan konsumsi pangan, capaiannya adalah berhasil membentuk 4 cluster optimum dengan masing-masing anggota pada cluster 1 sebanyak 22 provinsi, pada cluster 2 sebanyak 10 provinsi, pada cluster 3 sebanyak 1 provinsi, dan pada cluster 4 sebanyak 1 provinsi [15]. Asrul Asahri Muin juga melakukan penelitian tentang SOM yaitu implementasi SOM untuk klasifikasi penduduk dalam menentukan keputusan pembangunan daerah prioritas miskin di Kota Makassar. Pokok pembahasannya yaitu informasi pemetaan penduduk miskin berdasarkan karekteristik demografi, pendidikan, ketenagakerjaan, perumahan dan membantu pengambilan kebijakan dalam pemerataan program kemiskinan sehingga dapat melakukan pembangunan daerah secara prioritas [16], selain itu penelitian SOM untuk pengelompokkan jurusan dilakukan oleh Rusydi Umar dengan capaiannya yaitu, hasil cluster dapat digunakan sebagai rekomendasi pada calon peserta didik yang memiliki skill dan bakat serta minat yang dimiliki [7] .

2. METODOLOGI PENELITIAN

2.1 Metode Penelitian

Penelitian ini dilakukan di fakultas teknik informatika Universitas Budi Darma dengan mengumpulkan dataset nilai mahasiswa berdasarkan matakuliah yang digunakan dan melakukan *clustering* sesuai algoritma SOM.

Metode penelitian untuk pengumpulan data dan analisis terdiri dari.

a. Studi Literatur

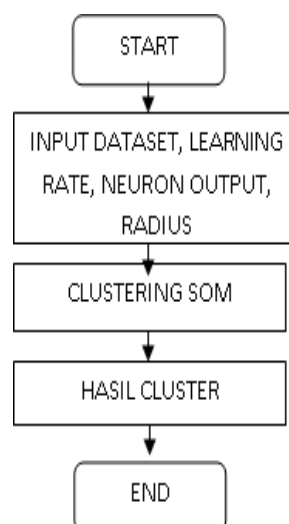
Pada tahap ini mencari literatur dengan rincian sebagai berikut :

1. Mencari referensi tentang teori dan pengetahuan yang membahas tentang bidang karya ilmiah (skripsi) dan *machine learning*.
2. Mencari refrensi mengenai algoritma *clustering*, *Self Organizing Maps*, dan dalam ruang lingkup *machine learning*.

b. Faktor yang perlu diperhatikan dalam desain sistem antara lain :

1. Dataset nilai-nilai matakuliah pendukung karya ilmiah *machine learning*
2. Variabel input yang digunakan dalam algoritma SOM.
3. Faktor penentu klaster-klaster / *neuron output*

Deklarasi penetapan hasil cluster, variabel input, dan bobot disesuaikan dengan arsitektur SOM Kohonen Maps. Untuk di awal ini dijelaskan hasil cluster dalam bentuk bidang karya ilmiah (skripsi) yang akan dikerjakan mahasiswa. Berikut gambaran alur penelitian yang dilakukan dalam cakupan input, Proses, dan output.



Gambar 1. Alur penelitian

2.2 Dataset, Variabel Input, Dan Bobot

Gambar 3 Alur penelitian adalah bagian-bagian yang telah di-set dalam penelitian ini. Dimulai dari bagian input yaitu input *dataset*, *Learning rate*, radius, dan variabel lainnya yang digunakan baik dalam analisa dan penerapan selanjutnya. Hal dirincikan sebagai berikut :

a. Input *Dataset*, merupakan vektor atau variabel input digunakan tersimpan dalam *dataset* adalah nilai-nilai matakuliah dari setiap mahasiswa sebagai pendukung *machine learning* yaitu :

1. Matakuliah Algoritma dan Pemrograman (MT1)
2. Matakuliah Statistik dan Probalitas (MT2)
3. Matakuliah Matrik dan Aljabar Linear (MT3)
4. Matakuliah Struktur Data (MT4)

Matakuliah-matakuliah di atas diinisialkan dengan MT dan nomor urutan yaitu MT1, MT2, MT3, dan MT4. Nilai-nilai matakuliah yang digunakan di atas terwujud dalam *range* nilai seperti pada tabel 1 di bawah. Namun secara tepatnya nilai tersebut diubah atau diambil yang terdapat pada kolom bobot, serta diminimalkan dalam pembulatan yaitu, A = 4, B = 3, C = 2, dan D = 1. Untuk E = 0 tidak digunakan karena tidak lulus.

Tabel 1. Range nilai akhir setiap matakuliah

Nilai Total	Nilai Huruf (NH)	Bobot	Keterangan
≥90	A	4	Lulus
≥ 80 - ≤89	A-	3,75	Lulus
≥ 76 - ≤79	B+	3,5	Lulus
≥ 70 - ≤75	B	3	Lulus
≥ 67 - ≤69	B-	2,75	Lulus
≥ 65 - ≤66	C+	2,5	Lulus
≥ 60 - ≤64	C	2	Lulus
≥ 55 - ≤59	C-	1,5	Lulus
≥ 50 - ≤54	D	1	Lulus
< 50	E	0	Tidak Lulus

b. Variabel Input

1. Menetapkan $\alpha(t)$ sebagai *learning rate* dan $\sigma_2(t)$ adalah koefisien fungsi aktivasi antara 0 sampai 1.
2. Menetapkan *epoch* (Ω) = 0, yaitu jumlah berapa kali sebuah data dimasukkan ke dalam jaringan untuk proses latih sebelum ukuran neighbour berkurang pada setiap iterasi.
3. *Learning rate* (α) = 0.6 akan berubah *learning rate* dapat mengikuti fungsi geometri $(t+1)=0.5\alpha(t)$. Hasil proses hanya dua cluster, maka *neighbor* dari J mengubah satu cluster dengan radius (R) = 1.
4. Bobot diinisialisasikan dalam bilangan acak dari distribusi seragam (*uniform*). Fungsi meratakan distribusi sampel dengan batas interval separuh terbuka. Dengan kata lain, untuk semua nilai dalam batas interval yang ditentukan dapat diambil bersamaan dengan variabel rendah dan variabel tinggi. Pembuatan batas interval untuk inisialisasi bobot ini dapat dilakukan dengan tool pengolahan data sains python, bentuk umumnya seperti berikut : $w = np.random.uniform(-1,0,1000)$

2.3 Clustering Menggunakan Self Organizing Map (Kohonen)

Algoritma SOM terdiri sebagai berikut [17], [18]:

- a. Inisialisasi vektor input $x_1, x_2, x_3, \dots, x_n$ dengan nilai random 0-1.
- b. Inisialisasi neuron output sebanyak y_1, y_2, \dots, y_n .
- c. Menentukan *weight* (bobot) neuron output dengan nilai antara x_{\min} dan x_{\max} .
- d. Mengulangi langkah 5 sampai 8 hingga tidak ada *update weight* (bobot) atau telah mencapai kondisi stop (*error* terkecil).
- e. Pemilihan acak salah satu data dari vektor input sebagai data *training*.
- f. Mencari jarak terdekat dari masing-masing *neuron output* ke data input menggunakan rumus *euclidian distance*. $D(j) = \sum_i (w_{ij} - x_i)^2$ $2D(j) = \sum_i (w_{ij} - x_i)^2$ (1)
- g. Untuk setiap bobot w_{ij} diperbaharui bobot tetangga menggunakan rumus dengan persamaan sebagai berikut $w_{ij}(t+1) = w_{ij}(t) + \alpha(t)[x_i - w_{ij}(t)]$ (2)
- h. Mengupdate bobot bias (*error*).
- i. Simpan bobot yang telah konvergen.

2.4 Hasil Cluster

Target hasil cluster terdiri dari dua bagian yaitu :

- a. Cluster pertama diinisialkan *cluster* 0 yang termasuk ke dalam karya ilmiah yang akan dikerjakan dalam konsep *machine learning*.
- b. Cluster kedua diinisialkan *cluster* 1 yang tidak termasuk dari karya ilmiah yang akan dikerjakan dalam konsep *machine learning*.

3. ANALISA DAN PEMBAHASAN

3.1 Analisa

Kegiatan melakukan pengujian dari objek *dataset* dan variabel input yang terukur sebagai data masukan. Data masukan ini ditargetkan untuk menghasilkan dua *cluster* yang menyatakan sanggup atau tidak mengkaji karya ilmiah dalam lingkup *machine learning*.

3.2 Sampel Data

Pengujian variabel input dalam vektornya terdiri dari 40 sampel data mahasiswa. Berikut adalah sampel data yang digunakan.

Tabel 2. Sampel data mahasiswa Universitas Budi Darma

No	Nama Mahasiswa	MT1	MT2	MT3	MT4
1	Alexander P. Manullang	2	4	3	3
2	Annisa Apriliani	3	3	3	3
3	Delima R Tambunan	3	2	2	3
4	Ewit Supriana Purba	2	3	3	4
5	Heber Sinurat	3	3	2	3
6	Jordan Purba	4	4	2	3
7	Kevin Yanto Sarumaha	3	3	2	4
8	Lastiar Pardede	3	2	3	3
9	Reka Wulandari	2	2	2	3
10	Ronal Marco	4	3	4	4
11	Rudianto Nehe	4	4	2	3
12	Sierman Telaumbanua	3	2	2	4
13	Theresia Mardiana Siregar	3	2	3	3
14	Vira Azura	2	2	2	3
15	Diffri Solihin Siregar	4	4	4	4
16	Afflina Simbolon	3	4	3	3
17	Ahmad Ismail	2	3	2	4
18	Aldi Umri Wirya Chandra	3	2	3	3
19	Alvina Novanda	2	3	2	3
20	Amanudin Harahap	4	2	4	4
21	Arif Budiman	2	4	2	4
22	Cici Rahmadani	3	2	2	4
23	Daniel D Saragih	3	2	3	2
24	Elvya Dani Yanty Sitepu	2	2	3	3
25	Fransiscus Situmorang	4	2	4	3
26	Galuh Kesuma Pramesti	4	3	3	4
27	Irsyan Syauqi Siregar	3	2	2	3
28	Lasmi Br Banurea	2	4	2	4
29	Maringan Yoseptian Sitanggang	3	3	2	3
30	Marta Zega	4	2	2	3
31	Mei Putri Siringo-ringo	4	4	4	4
32	Muhammad Salman Alfarizhi	3	4	2	3
33	Noki Cahya Putra Pasaribu	2	3	2	4
34	Nurhasanah	3	2	3	2
35	Rephael Pasaribu	2	2	3	3
36	Rindi Novia	2	2	3	3
37	Robi Permana	4	3	4	2
38	Susi Lampita Sinaga	3	2	2	4
39	Syahbalyan Ari Lubis	3	2	3	3
40	Trilis Wulandari	4	2	3	2

3.3 Algoritma SOM

Proses SOM sangat sederhana, asalkan variabel-variabel yang berkaitan untuk proses SOM tersebut benar dan jelas. Pada bagian ini untuk proses SOM sendiri dilakukan menggunakan fungsi pemrograman python. Berikut adalah langkah-langkah Algoritma SOM[19].

Langkah 0 Inisialisasi bobot: w_{ij}

set parameter tetangga

set learning rate α

Langkah 1 Ketika kondisi berhenti bernilai *false*, kerjakan Langkah 2

Langkah 2 Untuk setiap vektor input x , kerjakan Langkah 3-5

Langkah 3 Untuk setiap j hitung $D(j)=\sum_i(w_{ij}-x_i)^2$

Langkah 4 Tentukan J , sedemikian sehingga $D(J)$ minimum

Langkah 5 Pada setiap unit j dengan spesifikasi tetangga tertentu dari J , dan untuk setiap i maka $w_{ij}(\text{baru})=w_{ij}(\text{lama})+\alpha[x_i-w_{ij}(\text{lama})]$

Langkah 6 Perbarui learning rate

Langkah 7 Kurangi radius ketetanggaan pada waktu-waktu tertentu

Langkah 8 Tes kondisi berhenti

Dari algoritma di atas diwujudkan dalam kode program python untuk setiap langkah-langkahnya yang difungsikan sebagai prosedur SOM seperti pada gambar 2 di bawah berikut :

```
def KohonenSOM(x_unit,n_cluster,max_epoch,learning_rate=1,lr_decrease=0.5,radius=1):
    # Langkah 0
    w = np.random.uniform(size=(x_unit.shape[1],n_cluster))
    epoch = 0
    # Langkah 1
    while epoch < max_epoch:
        stop = True
        # Langkah 2
        for x in x_unit:
            # Langkah 3
            d = [sum((w[:,j] - x)**2) for j in range(w.shape[1])]
            # Langkah 4
            min_index = np.argmin(d)
            # Langkah 5
            left_index = min_index - radius
            right_index = min_index + radius
            # ketika ketanggaan j diluar batas index
            if left_index < 0: left_index = 0
            if right_index >= w.shape[1]: right_index = w.shape[1]-1
            temp_w = np.copy(w)
            for j in range(left_index,right_index+1):
                w[:,j] += learning_rate*(x - w[:,j])
            delta_w = w - temp_w
            if not np.array_equal(delta_w,np.zeros(delta_w.shape)):
                stop = False
        # Langkah 6
        learning_rate *= lr_decrease
        # Langkah 7
        if radius>0 and epoch%5 == 0:
            radius -= 1
        # Langkah 8
        if stop:
            break
        epoch += 1
    return w
```

Gambar 2. Prosedur SOM dalam bentuk kode Python

3.4 Pembahasan

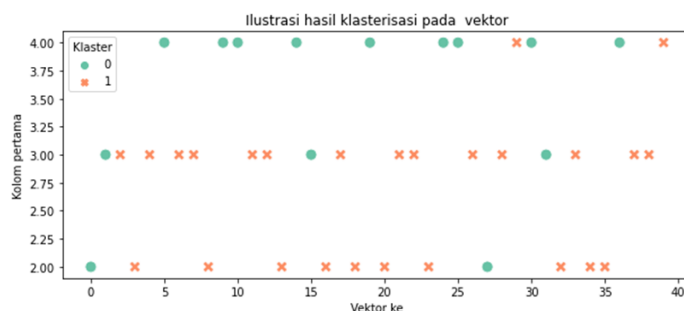
Proses ini sesuai algoritmanya diawali dari inisialisasi bobot secara distribusi *random uniform* dan diterangkan radius ketanggaan (*neighbour*) dilakukan dengan menggunakan *array* linear dari *cluster* unit, posisi radius berkurang setiap setelah 5x *epoch* dan kondisi berhenti ketika tidak ada perubahan bobot yang terjadi. Hasil *cluster* diketahui untuk yang tergabung *cluster* 0 sebanyak 14 mahasiswa dan untuk *cluster* 1 sebanyak 26 mahasiswa. Maka mahasiswa yang berada pada *cluster* 0 dianjurkan dan diarahkan untuk berencana/mengangkat topik karya ilmiah berbasis *machine learning*, sedangkan mahasiswa yang termasuk pada *cluster* 1 sangat tidak dianjurkan mengangkat topik karya ilmiah berbasis *machine learning*. Berikut adalah hasil *cluster* dari SOM.

Tabel 3. Hasil dua cluster SOM

No	Nama Mahasiswa	MT1	MT2	MT3	MT4	Cluster
1	Alexander P. Manullang	2	4	3	3	0
2	Annisa Apriliani	3	3	3	3	0

No	Nama Mahasiswa	MT1	MT2	MT3	MT4	Cluster
3	Delima R Tambunan	3	2	2	3	1
4	Ewit Supriana Purba	2	3	3	4	1
5	Heber Sinurat	3	3	2	3	1
6	Jordan Purba	4	4	2	3	0
7	Kevin Yanto Sarumaha	3	3	2	4	1
8	Lastiar Pardede	3	2	3	3	1
9	Reka Wulandari	2	2	2	3	1
10	Ronal Marco	4	3	4	4	0
11	Rudianto Nehe	4	4	2	3	0
12	Sierman Telaumbanua	3	2	2	4	1
13	Theresia Mardiana Siregar	3	2	3	3	1
14	Vira Azura	2	2	2	3	1
15	Diffri Solihin Siregar	4	4	4	4	0
16	Afflina Simbolon	3	4	3	3	0
17	Ahmad Ismail	2	3	2	4	1
18	Aldi Umri Wirya Chandra	3	2	3	3	1
19	Alvina Novanda	2	3	2	3	1
20	Amanudin Harahap	4	2	4	4	0
21	Arif Budiman	2	4	2	4	1
22	Cici Rahmadani	3	2	2	4	1
23	Daniel D Saragih	3	2	3	2	1
24	Elvya Dani Yanty Sitepu	2	2	3	3	1
25	Fransiscus Situmorang	4	2	4	3	0
26	Galuh Kesuma Pramesti	4	3	3	4	0
27	Irsyan Syauqi Siregar	3	2	2	3	1
28	Lasmi Br Banurea	2	4	2	4	0
29	Maringan Yoseptian Sitanggang	3	3	2	3	1
30	Marta Zega	4	2	2	3	1
31	Mei Putri Siringo-ringo	4	4	4	4	0
32	Muhammad Salman Alfarizhi	3	4	2	3	0
33	Noki Cahya Putra Pasaribu	2	3	2	4	1
34	Nurhasanah	3	2	3	2	1
35	Rephael Pasaribu	2	2	3	3	1
36	Rindi Novia	2	2	3	3	1
37	Robi Permana	4	3	4	2	0
38	Susi Lampita Sinaga	3	2	2	4	1
39	Syabhalyan Ari Lubis	3	2	3	3	1
40	Trilis Wulandari	4	2	3	2	1

Hasil dalam bentuk visual plot diagram dari cluster data yang diinputkan ditampilkan pada gambar 3.



Gambar 3. Visual cluster

Sebaran data pada diagram di atas ditampilkan arah sumbu x adalah vektor sampel data sebanyak 40 orang mahasiswa, sedangkan arah sumbu y interval nilai dari setiap matakuliah dengan interval 1 hingga 4, namun karena nilai data terkecil adalah 2, maka dimulai dari 2. Pada gambar terlihat yang terdata pada *cluster 0* berwarna hijau terletak pada posisi :

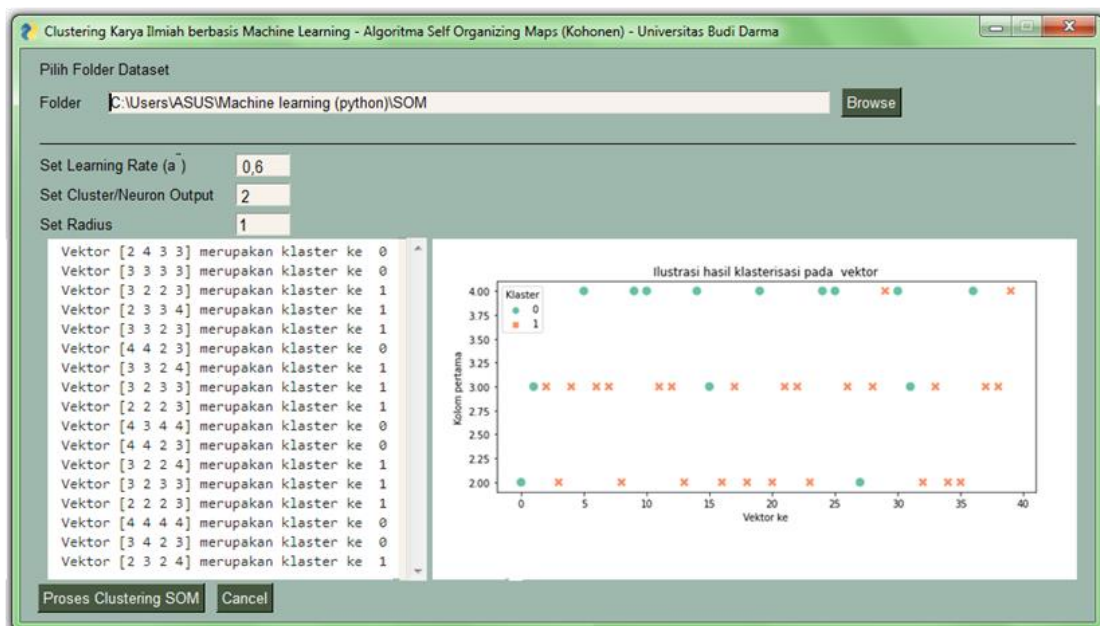
y = 2 sebanyak 2 mahasiswa

y = 3 sebanyak 3 mahasiswa

y = 4 sebanyak 9 mahasiswa

3.5 Implementasi

Penerapan aplikasi *clustering* karya ilmiah mahasiswa berbasis *machine learning* pada Universitas Budi Darma dapat dilakukan dengan sederhana namun memberikan hasil yang baik dan cepat untuk proses *clustering*. Pada tampilan aplikasi di gambar 4 dapat dilihat pertama memasukkan data uji yang tersimpan dengan data.csv pada tempat penyimpanan. Selanjutnya memasukkan nilai *learning rate* dengan 0.6, memasukkan jumlah *clustering* atau *neuron output* sebanyak 2 untuk menghasilkan *cluster 0* dan 1, dan selanjutnya nilai radius sebanyak 1. Untuk tampilan hasil pada bagian aplikasi terdapat dua bagian yaitu *multiline* yang menunjukkan hasil setiap klaster, serta di bagian kanan adalah bentuk *visual diagram clustering*-nya. Bentuk lain dari output ini yaitu saat tombol proses *clustering SOM* ditekan akan menampilkan hasil prosesnya dan membuat file hasil akhir tersebut dengan nama “hasil cluster SOM.txt”.



Gambar 4. Implementasi aplikasi *cluster SOM*

4. KESIMPULAN

Clustering karya ilmiah menggunakan *Self Organizing Maps* (Kohonen) memberikan hasil *clustering* yang baik jika nilai radius dapat difungsikan perubahan secara tepat. Selain itu nilai random untuk pembobotan juga berpengaruh dalam proses iterasi dengan *epoch* yang berkurang setiap 5 kali *epoch*. Bobot diinisialisasikan dalam bilangan acak dari distribusi seragam (*uniform*). Fungsi meratakan distribusi sampel dengan batas interval separuh terbuka. Dengan kata lain, untuk semua nilai dalam batas interval yang ditentukan dapat diambil bersamaan dengan nilai variabel *low* (min) dan variabel *high* (maks). Jumlah cluster yang diinginkan ada dua yaitu cluster 0 mampu mengerjakan karya ilmiah berbasis machine learning dan cluster 1 sebaliknya. Proses SOM terhadap 40 data mahasiswa dengan target sebanyak dua cluster dan hasilnya cluster 0 = 14 mahasiswa, cluster 1 = 26 mahasiswa. Hasilnya didapatkan dengan menaikkan radius = 1, yang mana sebelumnya pencapaian ini tidak berhasil jika radius = 0.

REFERENCES

- [1] A. S. Khazari, F. Marisa, and I. D. Wijaya, “Sistem Rekomendasi Penentuan Judul Skripsi Menggunakan Algoritma Decision Tree,” *J. Teknol. dan Manaj. Inform.*, vol. 3, no. 1, 2017, doi: 10.26905/jtmi.v3i1.1248.
- [2] I. Wahyudin, E. Tita Tosida, and F. Andria, “Teori dan Panduan Praktis Data Science dan Big Data,” *Quality*, no. March, pp. 1–6, 2020.
- [3] A. R. Sinaga and Y. Hasan, “Aplikasi Pendukung Keputusan Penentuan Dosen Pembimbing Skripsi S1 Teknik Informatika,” *Inf. dan Teknol. Ilm.*, pp. 182–187, 2015, [Online]. Available:

- https://s3.amazonaws.com/academia.edu.documents/40796550/31._Jurnal__ALEX.pdf?response-content-disposition=inline%3B+filename%3D31._Jurnal_ALEX.pdf&X-Amz-Algorithm=AWS4-HMAC-SHA256&X-Amz-Credential=ASIATUSBJ6BAPZF7ABG3%2F20200418%2Fus-east-1%2Faws3%2Faws4_r
- [4] R. R. Santoso, R. Megasari, and Y. A. Hambali, “Implementasi Metode Machine Learning Menggunakan Algoritma Evolving Artificial Neural Network Pada Kasus Prediksi Diagnosis Diabetes,” *JATIKOM (Jurnal Apl. dan Teor. Ilmu Komputer)*, vol. 3, no. 2, pp. 85–97, 2020.
- [5] Fahrizal, F. O. Reynaldi, and N. Hikmah, “Implementasi Machine Learning pada Sistem PETS Identification Menggunakan Python Berbasis Ubuntu,” *J. Inf. Syst. Informatics Comput.*, vol. 4, no. 1, pp. 86–91, 2020.
- [6] F. A. Bachtiar, I. K. Syahputra, and S. A. Wicaksono, “Perbandingan Algoritme Machine Learning untuk Memprediksi Pengambil Matakuliah,” *J. Teknol. Inf. dan Ilmu Komput.*, vol. 6, no. 5, p. 543, 2019, doi: 10.25126/jtiik.2019651755.
- [7] R. Umar, A. Fadlil, and R. R. Az-Zahra, “Self Organizing Maps(SOM) untuk Pengelompokkan Jurusan di SMK,” *Khazanah Inform. J. Ilmu Komput. dan Inform.*, vol. 4, no. 2, p. 131, 2018, doi: 10.23917/khif.v4i2.7044.
- [8] S. S. Nagari and L. Inayati, “Implementation of Clustering Using K-Means Method To Determine Nutritional Status,” *J. Biometrika dan Kependud.*, vol. 9, no. 1, p. 62, 2020, doi: 10.20473/jbk.v9i1.2020.62-68.
- [9] A. Roihan, P. A. Sunarya, and A. S. Rafika, “Pemanfaatan Machine Learning dalam Berbagai Bidang: Review paper,” *IJCIT (Indonesian J. Comput. Inf. Technol.)*, vol. 5, no. 1, pp. 75–82, 2020, doi: 10.31294/ijcit.v5i1.7951.
- [10] D. Rahmalia and T. Herlambang, “Application Kohonen Network and Fuzzy C Means for Clustering Airports Based on Frequency of Flight,” *Kinet. Game Technol. Inf. Syst. Comput. Network, Comput. Electron. Control*, vol. 3, no. 3, pp. 229–236, 2018, doi: 10.22219/kinetik.v3i3.608.
- [11] T. Honkela, “Description of Kohonen ’ s Self-Organizing Map,” pp. 1–7, 2018.
- [12] N. F. Arno and M. Ahsan, “Neural Network,” *Power Syst.*, vol. 28, pp. 75–159, 2017, doi: 10.1007/978-3-030-64777-3_4.
- [13] D. Sarkar, R. Bali, and T. Sharma, *Practical Machine Learning with Rust*. 2020.
- [14] L. S. Moonlight, F. Faizah, Y. Suprpto, and N. Pambudiyatno, “Comparison of Backpropagation and Kohonen Self Organising Map (KSOM) Methods in Face Image Recognition,” *J. Inf. Syst. Eng. Bus. Intell.*, vol. 7, no. 2, p. 149, 2021, doi: 10.20473/jisebi.7.2.149-161.
- [15] M. D. Kartikasari, “Self-Organizing Map Menggunakan Davies-Bouldin Index dalam Pengelompokan Wilayah Indonesia Berdasarkan Konsumsi Pangan,” *Jambura J. Math.*, vol. 3, no. 2, pp. 187–196, 2021, doi: 10.34312/jjom.v3i2.10942.
- [16] A. A. Muin, “IMPLEMENTASI SELF ORGANIZING MAPS (SOM) KLASIFIKASI PENDUDUK UNTUK MENENTUKAN KEPUTUSAN PEMBANGUNAN DAERAH PRIORITAS MISKIN (Studi Kasus Kota Makassar) Berdasarkan dirumuskan , apakah mengklasifikasi latar metode belakang SOM berdasar dapat dapat ting,” pp. 1–8, 2018.
- [17] G. Munawar, “Implementasi Algoritma Self Organizing Map (SOM) untuk Clustering Mahasiswa pada Matakuliah Proyek (Studi Kasus : JTK POLBAN),” *Pros. 6th Ind. Res. Work. Natl. Semin.*, pp. 66–78, 2015.
- [18] E. B. Perkasa, B. Wijaya, and J. Jamhari, “Pemodelan Jaringan Kohonen Untuk Pengenalan Nama Asteroid Kohonen Network Modeling for Asteroid Name Recognition,” vol. 10, no. 2, pp. 127–138, 2020.
- [19] R. N. R. Wijaya, “Kohonen Self-Organizing Maps,” pp. 3–5, [Online]. Available: <https://github.com/RinRoya/Kohonen-Self-Organizing-Maps/>.