

## **KETEPATAN KLASIFIKASI PENERIMA BEASISWA STIMIK STIKOM BALI DENGAN HYBRID SELF ORGANIZING MAPS DAN ALGORITMA K-MEANS**

**Ni Putu Nanik Hendayanti<sup>1</sup>, Gusti Ayu Made Arna Putri<sup>2</sup>, Maulida Nurhidayati<sup>3</sup>**

STIMIK STIKOM Bali<sup>1,2</sup>

Institut Agama Islam Negeri Ponorogo<sup>3</sup>

nanik@stikom-bali.ac.id<sup>1</sup>, gustiayu.arna@gmail.com<sup>2</sup>, nurhidayatimaulida@gmail.com<sup>3</sup>

### **Abstrak**

Data mining adalah penemuan informasi baru dengan mencari pola atau aturan tertentu dari sejumlah data yang sangat besar. Salah satu teknik yang dikenal dalam data mining yaitu clustering. Pengertian clustering dalam data mining adalah pengelompokan sejumlah data atau objek ke dalam cluster (group) sehingga setiap anggota dari cluster tersebut akan berisi data yang semirip mungkin dan berbeda dengan objek dalam cluster yang lain. Salah satu metode klasifikasi atau clustering adalah Self Organizing Maps (SOM). SOM merupakan metode artificial neural network yang digunakan untuk mengelompokkan (clustering) data berdasarkan karakteristik/fitur-fitur data. Metode pengelompokan yang menggunakan konsep jarak dan memiliki karakteristik yang hampir sama dengan SOM yaitu algoritma K-means. Penelitian ini mengembangkan suatu metode yang merupakan hybrid dari metode SOM dan algoritma K-means (selanjutnya disebut hybrid SOM dan K-means) yang digunakan untuk menentukan ketepatan suatu klasifikasi. Data yang digunakan dalam penelitian ini adalah data penerimaan beasiswa di STMIK STIKOM Bali pada tahun 2017. Hasil penelitian menunjukkan bahwa hasil klasifikasi yang diperoleh dengan algoritma K-means dan metode SOM memberikan hasil yang sama. Hybrid metode SOM dan K-means menggunakan nilai pusat dari hasil SOM untuk selanjutnya dilakukan pengelompokan dengan menggunakan algoritma K-means. Nilai pusat yang dimiliki metode SOM dan algoritma K-means sama sehingga hasil klasifikasi yang diperoleh dari hybrid metode SOM dan K-means juga memberikan hasil yang sama dengan metode SOM dan K-means. Akibatnya ketepatan klasifikasi untuk ketiga metode yang digunakan adalah sama yaitu 54,45%.

**Kata Kunci :** *Data Mining, SOM, Algoritma K-Means, hybrid SOM dan K-means*

---

### **I. PENDAHULUAN**

*Data mining* sering juga disebut sebagai *knowledge discovery in database* (KDD). KDD merupakan kegiatan yang meliputi pengumpulan, pemakaian data, historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar, sehingga istilah *pattern recognition* sekarang jarang digunakan karena sudah termasuk bagian dari *Data Mining*. Salah satu metode yang sering dipakai dalam literatur *data mining* yaitu *clustering* [1].

*Clustering* (Analisis klaster) bertujuan untuk mengalokasikan sekelompok individu pada suatu kelompok-kelompok yang saling bebas sehingga individu-

individu di dalam satu kelompok yang sama mirip satu sama lain, sedangkan individu-individu di dalam kelompok yang berbeda tidak mirip.

Salah satu metode klasifikasi atau *clustering* adalah *Self Organizing Maps* (SOM). SOM merupakan metode *artificial neural network* yang digunakan untuk mengelompokkan (*clustering*) data berdasarkan karakteristik/fitur-fitur data. Menurut [2] SOM merupakan salah satu bentuk topologi dari *Unsupervised Artificial Neural Network* (*Unsupervised ANN*) dimana dalam proses trainingnya tidak memerlukan pengawasan (target Output). Metode ini memungkinkan untuk menggambarkan data multidimensi ke dalam dimensi yang lebih kecil, biasanya satu atau dua dimensi.

---

Proses penyederhanaan ini dilakukan dengan mengurangi vektor yang menghubungkan masing-masing node. Cara ini disebut juga dengan *Vektor Quantization*. Teknik yang dipakai dalam metode SOM dilakukan dengan membuat jaringan yang menyimpan informasi dalam bentuk hubungan node dengan training set yang ditentukan.

Beberapa penelitian yang menggunakan SOM dalam pengelompokan diantaranya: [3] *Clustering of the Self-Organizing Map*, [4] *Principal Component Analysis and Self Organizing Map for Visualizing and Classifying Fire Risk in Forest Regions*, [5] *Classification of Indian power Coals Using K-means Clustering and Self Organizing Map Neural Network*.

Untuk mengatasi kekurangan visualisasi dari pengelompokan SOM, dibutuhkan metode pengelompokan yang menggunakan konsep jarak dan memiliki karakteristik yang hampir sama dengan SOM yaitu algoritma K-means. Algoritma *K-means* dipilih karena algoritma ini menggunakan ukuran kemiripan untuk mengelompokkan obyek. Kemiringan ini dapat diterjemahkan dalam konsep jarak. Menurut [1] Algoritma *clustering K-means* dapat diringkas dengan memilih jarak terdekat dengan pusat dilanjutkan menghitung pusat baru berdasarkan hasil pengelompokan. Oleh karena penggabungan metode SOM dan algoritma *K-means* (*hybrid SOM* dan *K-means*) dalam menentukan ketepatan klasifikasi sampai saat ini masih terbatas maka perlu dilakukan penelitian lebih lanjut tentang metode tersebut. Dalam penelitian ini, metode *hybrid SOM* dan *K-means* akan digunakan untuk menentukan ketepatan klasifikasi pada data penerimaan beasiswa STMIK STIKOM Bali tahun 2017.

Beasiswa adalah pemberian berupa bantuan keuangan yang diberikan kepada perorangan yang bertujuan untuk digunakan demi keberlangsungan pendidikan yang ditempuh [6]. Pemberian beasiswa merupakan program kerja yang ada di setiap universitas atau perguruan tinggi. Program beasiswa diadakan untuk meringankan beban mahasiswa dalam menempuh masa studi kuliah khususnya dalam masalah biaya

Setiap lembaga pendidikan banyak sekali beasiswa yang ditawarkan, mulai dari beasiswa berprestasi sampai beasiswa kurang mampu, salah satunya STMIK STIKOM Bali. STMIK STIKOM Bali telah mengadakan program beasiswa, antara lain beasiswa PPA (Peningkatan Prestasi Akademik), beasiswa BBM (Bantuan Belajar Mahasiswa) dan

Bidik Misi, serta beasiswa yang dibiaya oleh yayasan itu sendiri. Proses seleksi penerimaan beasiswa dilakukan secara manual yaitu dengan menginputkan satu persatu data mahasiswa ke dalam file *spreadsheet* kemudian melakukan *sorting* data mahasiswa seringkali menimbulkan beberapa permasalahan, antara lain membutuhkan waktu yang lama dan ketelitian yang tinggi. Selain itu, transparansi serta ketidakjelasan metodologi yang digunakan dalam proses komputasi penerimaan beasiswa juga menjadi salah satu permasalahan.

Dengan mempertimbangkan hal tersebut maka diperlukan suatu sistem yang dapat membantu dalam proses pengambilan suatu keputusan siapa saja mahasiswa yang direkomendasikan menerima beasiswa berdasarkan kriteria-kriteria yang telah ditentukan secara cepat dan tepat sasaran, sehingga tidak semua pendaftar akan menerima beasiswa, hanya yang memenuhi kriteria yang telah ditentukan yang dapat menerima beasiswa tersebut. Metode yang digunakan dalam pengambilan keputusan untuk seleksi penerimaan beasiswa adalah *hybrid SOM* dan *K-means*.

#### A. Self Organizing Maps (SOM)

Kohonen *Self Organizing Feature Maps*, disingkat SOFM atau lebih terkenal dengan istilah SOM ditemukan dan dikembangkan oleh Teuvo Kohonen, seorang profesor di Academy of Finland. Metode ini memungkinkan untuk menggambarkan data multidimensi ke dalam dimensi yang lebih kecil, biasanya satu atau dua dimensi. Proses penyederhanaan ini dilakukan dengan mengurangi Vektor Quantization.

Salah satu hal yang menarik dalam metode SOM adalah kemampuannya untuk belajar secara mandiri (*unsupervised learning*). Pada metode belajar secara mandiri, network akan belajar tanpa adanya target terlebih dahulu. Hal ini berbeda dengan beberapa metode neural network yang lain seperti back propagation, perceptron, dan sebagainya yang memerlukan adanya target saat proses learning dilaksanakan [4].

#### B. Algoritma Pembelajaran SOM

Menurut [4] SOM tidak memerlukan target output. Training pada SOM dilakukan dengan langkah-langkah berikut ini:

1. Inisialisasi *weight* pada masing-masing node. Sebelum training dimulai, *weight* masing-masing node diberikan inisialisasi terlebih

dahulu. Biasanya diset dengan nilai random yang kecil.

$$W_{ij} = \frac{\min P_j + \max P_j}{2} \quad (1)$$

2. Sebuah vektor dipilih secara random.
3. Setiap vektor input terpilih dihitung jarak dengan *weight* dari *node* dan dipilih jarak yang minimum sebagai *Best Matching Unit* (BMU). Untuk menentukan BMU, salah satu metode yang dipakai adalah dengan menghitung semua *node* dan menghitung jarak *Euclidean* antara *weight* masing-masing *node* dengan vektor input. *Node* yang mempunyai vektor paling mendekati vektor input, maka ditentukan sebagai BMU. Persamaan mencari jarak *Euclidean* adalah:

$$Dist = \sum_{i=0}^{i=n} (V_i + W_i)^2 \quad (2)$$

4. Menentukan *node* tetangga BMU

### C. Analisis K-means

*K-means* termasuk dalam *partitioning clustering* yaitu setiap data harus masuk dalam *cluster* tertentu dan memungkinkan bagi setiap data yang termasuk dalam *cluster* tertentu pada suatu tahapan proses, pada tahapam berikutnya berpindah ke *cluster* yang lain. *K-means* memisahkan data ke k daerah bagian yang terpisah, dimana k adalah bilangan integer positif [7]. Algoritma *K-means* sangat terkenal karena kemudahan dan kemampuannya untuk mengklasifikasi data besar dan outlier dengan sangat cepat *k-means* tersebut menggunakan ukuran kemiringan untuk mengelompokkan obyek. Berikut adalah langkah-langkah algoritma *k-means* [5]:

1. Penentuan Pusat Cluster Awal  
Dalam menentukan n buah pusat *cluster* awal dilakukan pembangkitan bilangan random yang mempresentasikan urutan data input.
2. Perhitungan jarak dengan Pusat Cluster  
Untuk mengukur jarak antara data dengan pusat *cluster* digunakan Euclidian distance.
3. Pengelompokkan Data  
Jarak hasil perhitungan akan dilakukan perbandingan dan dipilih jarak terdekat antara data dengan pusat *cluster*, jarak ini menunjukkan bahwa data tersebut berada dalam satu kelompok dengan pusat *cluster* terdekat.
4. Penentuan Pusat Cluster Baru

Untuk mendapatkan pusat *cluster* baru bisa dihitung dari rata-rata nilai anggota *cluster* dan pusat *cluster*. Algoritma penentuan pusat *cluster* baru.

- a. Cari jumlah anggota tiap *cluster*
- b. Hitung pusat baru dengan rumus

$$Z_c = \frac{1}{n} \sum_{i=1}^n X_i, \quad i = 1, 2, \dots, n \quad (3)$$

## II. METODOLOGI

Data yang digunakan dalam penelitian ini adalah sekunder yang diperoleh dari bagian BAAK STMIK-STIKOM Bali yang berupa data beasiswa yang diterima mahasiswa yang dikumpulkan oleh mahasiswa pada saat pendaftaran beasiswa. Adapun variabel yang digunakan dalam penelitian ini sebagai berikut:

- Y = Klasifikasi penerimaan beasiswa
- $X_1$  = Nilai Indeks Prestasi Kumulatif (IPK) mahasiswa
- $X_2$  = Jumlah penghasilan Orang Tua per bulan
- $X_3$  = Semester  
Semester yang dimaksudkan disini adalah mahasiswa pada saat mengajukan beasiswa berada pada semester berapa
- $X_4$  = Jumlah tanggungan keluarga

Tahapan penelitian yang digunakan dalam penelitian ini adalah sebagai berikut

1. Pengumpulan data empiris pada bagian BAAK STMIK-STIKOM Bali yang berupa data pengajuan beasiswa oleh mahasiswa STIKOM Bali.
2. Deskripsi data masing-masing kelompok
3. Klasifikasi berdasarkan algoritma K-means
4. Klasifikasi berdasarkan metode Self-Organizing Maps (SOM).
5. Klasifikasi berdasarkan metode *hybrid* SOM dan K-means
6. Menghitung ketepatan klasifikasi yang dihasilkan dari masing-masing metode untuk mengetahui metode yang paling baik.

## III. HASIL dan PEMBAHASAN

### 1. Deskripsi Data

Sebelum analisis lebih lanjut, terlebih dahulu dilakukan deskripsi dari masing-masing kelompok untuk mempelajari karakteristik dari masing-masing

kelompok tersebut. Secara umum, kelompok 1 adalah kelompok penerima beasiswa PPA yang menerima beasiswa berdasarkan kemampuan akademiknya. Sehingga pada beasiswa PPA ini banyak merupakan mahasiswa yang secara akademik berprestasi. Untuk kelompok 2 adalah kelompok penerima BBP-PPA yang menerima beasiswa berdasarkan faktor yang lain selain dari pertimbangan IPK. Karakteristik masing-masing kelompok ditunjukkan pada Tabel 1.

**Tabel 1. Mean Variabel X1 dan X2**

Variable	Y	N	Mean
X1	1	85	3,692
	2	106	3,4658
X2	1	85	4,645,296
	2	106	4,292,049

Berdasarkan Tabel 1 diketahui bahwa mahasiswa yang memperoleh beasiswa PPA sebanyak 85 sedangkan mahasiswa yang memperoleh beasiswa BBP-PPA sebanyak 106. Kelompok mahasiswa penerima beasiswa PPA memiliki rata-rata IPK lebih tinggi dibandingkan dengan penerima beasiswa BBP-PPA. Hal ini sesuai dengan jenis pembagian beasiswanya dimana beasiswa PPA diperuntukkan bagi mahasiswa yang memiliki kemampuan akademik lebih tinggi. Untuk penerima beasiswa BBP-PPA diperuntukkan bagi mahasiswa yang secara finansial kurang (kurang mampu) sehingga rata-rata gaji orang tua untuk kelompok penerima beasiswa BBP-PPA lebih kecil dibandingkan dengan penerima PPA. Hal tersebut dapat dilihat pada hasil pengujian mean dari X2 yang menunjukkan bahwa penghasilan orang tua pada kelompok penerima beasiswa BBP-PPA lebih kecil dibandingkan dengan beasiswa PPA.

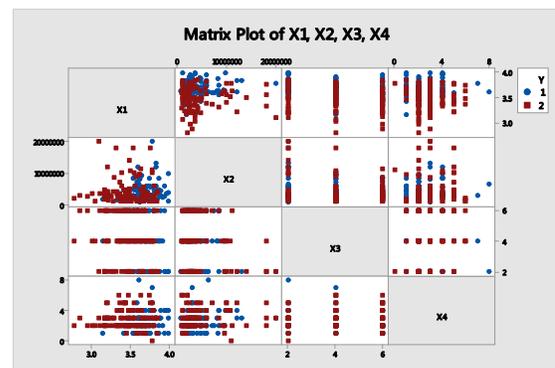
Variabel X3 dan X4 tidak dilakukan pencarian mean karena kurang dapat mendeskripsikan data yang sesungguhnya sehingga pada variabel X3 dan X4 digunakan ukuran modus untuk mengetahui bagaimana kecenderungan dari data tersebut dan dapat mewakili kondisi sebenarnya dari data. Hasil perhitungan modus dari variabel X3 dan X4 ditunjukkan pada Tabel 2.

**Tabel 2. Modus Variabel X1 dan X2**

Variable	Y	N	Mode
X3	1	85	4
	2	106	4

X4	1	85	2
	2	106	2

Tabel 2 menunjukkan modus untuk masing-masing kelompok untuk variabel semester dan jumlah tanggungan orang tua. Dari hasil tersebut diketahui bahwa modus untuk masing-masing kelompok sama. Sebagian besar mahasiswa yang menerima beasiswa adalah semester 4 dengan banyaknya tanggungan orang tua adalah 2.



**Gambar 1 Scatterplot Penerima Beasiswa**

Gambar 1 menunjukkan pengelompokan data berdasarkan pada masing-masing variabel. Berdasarkan Gambar 1 diketahui bahwa tidak terdapat pengelompokan secara jelas untuk masing-masing penerima beasiswa. Sehingga tidak dapat ditentukan secara pasti kelompok penerima beasiswa berdasarkan data yang dimiliki. Terlihat dari tercampurnya warna merah dan biru pada masing-masing variabel.

## 2. Klasifikasi Berdasarkan Algoritma K-Means

Deskripsi data yang dijabarkan sebelumnya selanjutnya digunakan dalam menentukan apakah kelompok yang sudah terbentuk merupakan kelompok penerima beasiswa PPA atau BBP-PPA. Hal ini dikarenakan penentuan kriteria kelompok pada algoritma *K-means* didasarkan pada kedekatan antar observasi sehingga pengetahuan tentang deskripsi awal dari masing-masing kelompok dapat membantu menentukan secara jelas kelompok yang telah terbentuk dengan algoritma *K-means* tersebut.

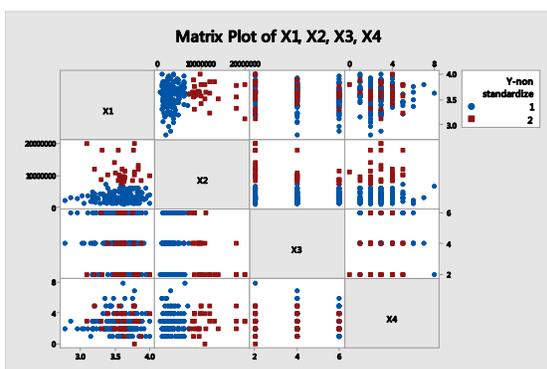
Penentuan kelompok dengan menggunakan algoritma *K-means* dilakukan dengan menggunakan *software* minitab 17. Hasil pengelompokan yang diperoleh ditunjukkan pada Tabel 3.

**Tabel 3 Hasil Kluster Algoritma K-means**

Variable	Y-ns	N	Mean
X1	1	161	3,5588
	2	30	3,6077
X2	1	161	3149090
	2	30	11426800

Tabel 3 menunjukkan hasil pengelompokan berdasarkan algoritma *K-means*. Untuk menentukan apakah kluster 1 adalah PPA atau BBP-PPA perlu dilakukan deskripsi data untuk masing-masing kluster. Seperti yang dijelaskan sebelumnya bahwa penerima beasiswa PPA memiliki nilai rata-rata IPK lebih tinggi dibandingkan dengan penerima BBP-PPA. Selain itu, penerima BBP-PPA adalah mahasiswa yang kurang mampu sehingga rata-rata penghasilan orang tua lebih kecil dibandingkan rata-rata penghasilan orang tua penerima PPA. Dari hasil yang diperoleh, diketahui bahwa kluster 1 adalah penerima beasiswa BBP-PPA karena karakteristik dari datanya memiliki IPK dan penghasilan orang tua lebih rendah dibandingkan dengan kluster 2. Jadi kluster 1 mewakili penerima beasiswa BBP-PPA sedangkan kluster 2 adalah penerima beasiswa PPA.

Gambar 2 menunjukkan hasil pengelompokan data berdasarkan pada masing-masing variabel. Berdasarkan gambar tersebut diketahui bahwa mahasiswa dapat terkluster secara terpisah terlihat dari tidak adanya yang *overlapping* dengan mahasiswa penerima beasiswa yang lain. Hal ini dapat dilihat dari kedua warna yang terpisah dan tidak tercampur.



Gambar 2 Scatterplot Penerima Beasiswa Berdasarkan Algoritma *K-means*

### 3. Klasifikasi Berdasarkan Metode SOM

Penentuan anggota kluster berdasarkan metode SOM hampir sama dengan algoritma *K-means*. Pada penelitian ini banyaknya kluster yang digunakan adalah 2 dengan

banyaknya *epoch*(iterasi) sebanyak 500 kali. Syntax yang digunakan untuk mendapatkan kluster dengan metode SOM ditunjukkan pada program dibawah ini:

```
data=xlsread('E:\data
beasiswa.xlsx',1,'A:E');
P = data(:,1:4)';
R=minmax(P)
net = newc(R,2);
bobot_input=net.IW{1,1}
bobot_bias_input=net.b{1,1}
net.trainParam.epochs=500;
net.trainParam.goal=0.001;
net=train(net,P)
bobot_akhir_input=net.IW{1,1}
bobot_akhir_bias_input=net.b{1,1}
Y = sim(net,P)
Yc = vec2ind(Y)
```

Bobot awal ditentukan berdasarkan default yang ada pada Matlab dengan menggunakan perintah *net.IW{1,1}* dimana penentuan bobotnya didasarkan pada titik tengah dari minimum dan maksimum dari input. Setelah bobot awal ditentukan, selanjutnya ditentukan bobot bias awal berdasarkan default yang ada pada Matlab dengan perintah *net.b{1,1}* dimana penentuan bobot biasnya diformulasikan dengan

$$\text{format } b = e^{1 - \ln\left(\frac{1}{\text{jumlah kluster}}\right)} = b = e^{1 - \ln\left(\frac{1}{2}\right)} = 5,4366.$$

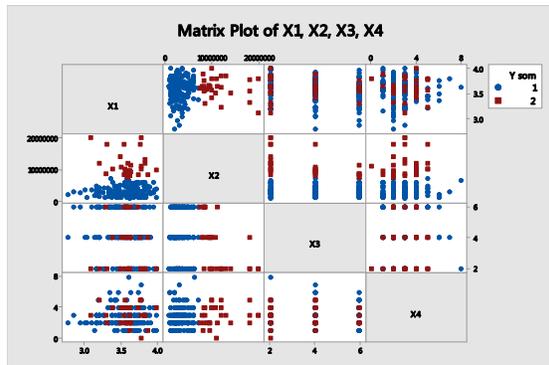
Dilanjutkan dengan melatih jaringan dengan *epoch* 500 serta toleransi error (untuk memastikan bahwa hasil yang diperoleh telah konvergen) sebanyak 0,001. Bobot awal dan bobot akhir yang diperoleh dengan metode ini tidak sama. Hal ini dikarenakan adanya pelatihan yang dilakukan hingga diperoleh hasil yang konvergen. Setelah konvergen, dilakukan simulasi dan diperoleh hasil klasifikasi seperti ditunjukkan pada Tabel 4.

Tabel 4 Hasil Klasifikasi Metode SOM

Variable	Y som	N	Mean
X1	1	161	3,5588
	2	30	3,6077
X2	1	161	3149090
	2	30	11426800

Hasil yang diperoleh pada Tabel 4 menunjukkan bahwa data penerima beasiswa terkluster menjadi 2 kluster dengan kluster 1 terdiri dari 161 anggota dan kluster 2 terdiri dari 30 anggota. Berdasarkan Tabel 4 tersebut juga diketahui bahwa rata-rata IPK kluster 1 kurang dari kluster 2 dan rata-rata penghasilan orang tua pada kluster 1 juga kurang dari kluster 2, sehingga dapat disimpulkan bahwa

kluster 1 adalah mahasiswa penerima beasiswa BBP-PPA sedangkan kluster 2 adalah penerima beasiswa PPA.



Gambar 3 Scatterplot Penerima Beasiswa Berdasarkan Metode SOM

Gambar 3 menunjukkan pengelompokan mahasiswa penerima beasiswa berdasarkan pada masing-masing variabel. Seperti halnya pada Gambar 2, Gambar 3 menunjukkan hasil kluster yang dapat memisahkan masing-masing mahasiswa berdasarkan variabel yang digunakan.

#### 4. Klasifikasi Dengan Hybrid Som Dan K-Means

Penentuan anggota kluster dengan hybrid SOM dan K-means hampir sama dengan kedua metode sebelumnya. Perbedaannya adalah pada metode ini menggabungkan dua metode klasifikasi untuk mendapatkan hasil yang lebih baik dan sesuai dengan data yang dimiliki.

Prinsip kerja dari metode ini adalah menentukan anggota kluster dengan metode SOM dan menghitung nilai pusat dari masing-masing kluster. Nilai pusat yang diperoleh dari masing-masing kluster selanjutnya digunakan untuk menentukan keanggotaan kluster pada algoritma K-means. Hasil kluster dengan menggunakan metode ini ditunjukkan pada Tabel 5.

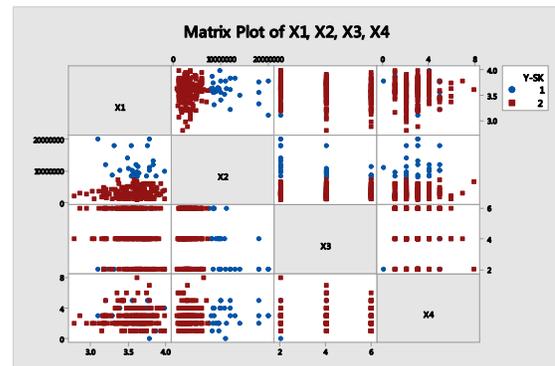
Tabel 5. Hasil Klasifikasi Metode SOM-Kmeans

Variable	Y-SK	N	Mean
X1	1	30	3,6077
	2	161	3,5588
X2	1	30	11426800
	2	161	3149090

Hasil yang diperoleh pada Tabel 5 menunjukkan bahwa data penerima beasiswa terkluster

menjadi 2 kluster dengan kluster 1 terdiri dari 30 anggota dan kluster 2 terdiri dari 161 anggota. Berdasarkan Tabel 5 tersebut juga diketahui bahwa rata-rata IPK kluster 1 lebih dari kluster 2 dan rata-rata penghasilan orang tua pada kluster 1 juga lebih dari kluster 2, sehingga dapat disimpulkan bahwa kluster 1 adalah mahasiswa penerima beasiswa PPA sedangkan kluster 2 adalah penerima beasiswa BBP-PPA.

Gambar 4 menunjukkan hasil pengelompokan mahasiswa penerima beasiswa berdasarkan pada masing-masing variabel. Seperti halnya pada Gambar 2 dan Gambar 3, Gambar 4 juga menunjukkan hasil kluster yang dapat memisahkan masing-masing mahasiswa berdasarkan variabel yang digunakan. Tidak terlihat ada *overlapping* pada masing-masing variabel.



Gambar 4 Scatterplot Penerima Beasiswa Berdasarkan hybrid SOM dan K-means

#### 5. Ketepatan Klasifikasi

Tahapan selanjutnya membandingkan hasil klasifikasi pada ketiga metode yang digunakan. Dengan membandingkan hasil kluster yang dimiliki dengan data kluster yang sesungguhnya, akan diperoleh ketepatan klasifikasi pada masing-masing metode yang digunakan. Ketepatan klasifikasi untuk masing-masing metode ditunjukkan pada Tabel 6.

Tabel 6. Ketepatan Klasifikasi

No	Metode	Ketepatan klasifikasi
1	Algoritma K-means	54,45
2	Metode SOM	54,45
3	Hybrid SOM dan K-means	54,45

Tabel 6 menunjukkan ketepatan klasifikasi pada masing-masing metode. Berdasarkan hasil tersebut diketahui bahwa ketiga metode yang digunakan memberikan hasil ketepatan klasifikasi yang sama yaitu 54,45%. Artinya variabel IPK, Penghasilan orang tua, semester, dan jumlah tanggungan keluarga dapat membagi penerima beasiswa dengan tingkat ketepatan 54,45% sisanya sebesar 45,55% adalah karena pengaruh variabel yang lain yang belum masuk dalam model.

Hasil yang diperoleh pada Tabel 6 juga menunjukkan bahwa pada kasus ini, tidak ada metode yang memberikan hasil yang paling baik. Ketiga metode memberikan hasil yang sama. Hal ini dapat terjadi karena nilai pusat yang diperoleh pada metode SOM, algoritma *K-means*, dan *hybrid* SOM dan *K-means* sama.

#### IV. KESIMPULAN

Algoritma *K-mean* dan metode SOM adalah beberapa metode pada analisis data yang bertujuan untuk klasifikasi terhadap obyek berdasarkan variabelnya. Prinsip kerja dari kedua metode ini sama yaitu dengan menentukan nilai pusat untuk kemudian membentuk klaster berdasarkan nilai pusat tersebut. Hasil klasifikasi pada kedua metode ini bisa sama dan bisa jadi berbeda tergantung pada nilai pusat yang digunakan. Ketika nilai yang digunakan sebagai pusat adalah sama, maka hasil klasifikasi yang dihasilkan juga akan sama. Akibatnya ketepatan klasifikasi yang dihasilkan juga akan sama. Jika nilai pusat yang digunakan pada algoritma *K-means* dan metode SOM tidak sama, dapat digunakan metode *hybrid* SOM dan *K-means* untuk mendapatkan hasil ketepatan yang lebih baik.

Pada penelitian ini diperoleh ketepatan klasifikasi yang sama untuk ketiga metode yang digunakan. Hal ini bisa dikarenakan nilai pusat yang dihasilkan sama. Akan tetapi pada penelitian ini belum secara jelas menampilkan nilai pusat yang digunakan pada masing-masing metode sehingga pada penelitian selanjutnya disarankan untuk memunculkan nilai pusat untuk masing-masing metode. Tujuannya adalah untuk mendeteksi apakah metode yang digunakan menghasilkan hasil ketepatan yang sama atau tidak. Selain itu juga dapat digunakan untuk menunjukkan bahwa dengan nilai pusat yang berbeda akan memberikan hasil yang berbeda.

Pada penelitian ini, data yang digunakan masih menggunakan 2 klaster sehingga pada penelitian

selanjutnya dapat dikembangkan lagi metode *hybrid* SOM dan *K-means* dengan menggunakan klaster lebih dari 2 atau menggunakan 2 klaster dengan nilai pusat yang berbeda.

#### V. KEPUSTAKAAN / REFERENSI

- [1] Santosa, B. (2007). Data Mining Teknik Pemanfaatan Data untuk keperluan Bisnis. Graha Ilmu. Yogyakarta.
- [2] Kohonen, T., (2001). *Self Organizing maps*. 3 penyunt. Berlin Heidelberg: Springer-Verlag.
- [3] Vesanto, J., (1999). SOM-based data Visualization Methods. *In Intelligent Data Analysis*, 3(2), pp. 111-126
- [4] Annas, S., Kanai, T. & S, K., (2007). Principal Component Analysis (PCA) and Self Organizing Map (SOM) for Visualizing and Classifying Fire Riks in Forest Region. *Agricultural Information Research*, 16(2), pp. 44-51.
- [5] Pandit, Y. P., Badhe, Y. P. & Sharma, B., (2011). Classification of Indian Power Coals using K-means Clustering and Self Organizing Map Neural Network. *Fuel*, (90), pp. 339-347.
- [6] Prasetyo, E. (2012). *Data Mining: Konsep dan Aplikasi Menggunakan Matlab*. Yogyakarta: ANDI Yogyakarta.
- [7] Johnson, R.A & Wichern, D.W. (2007). *Applied Multivariate Analysis*, 6th edition. New Jersey: Pearson Prentice Hall.

*Ketepatan Klasifikasi Penerima Beasiswa STMIK STIKOM Bali Dengan Hybrid Self Organizing Maps dan Algoritma K-Means*

**Ni Putu Nanik Hendayanti, Gusti Ayu Made Arna Putri, Maulida Nurhidayati**

---