

Penerapan Data Mining dalam pengelompokan Bibit Padi Unggul Berdasarkan Minat Beli Konsumen Pada PT.Sang Hyang Seri Regional IV Deli Serdang dengan Menggunakan Metode Clustering Algoritma K-Means

Suardi Yakub

Program Studi Sistem Informasi, STMIK Triguna Dharma

Article Info

Article history:

Received May 13th, 2018

Revised July 22th, 2018

Accepted Aug 05th, 2018

Keyword:

Bibit Unggul

Data Mining

K-Means

ABSTRACT

Perbedaan antara benih unggul dengan benih padi tidak unggul terletak pada proses sertifikasi, dimana benih bibit unggul di proses dan di pelihara sedemikian rupa sehingga tingkat kemurnian varietas dapat terpelihara dan memenuhi standar mutu benih yang ditetapkan serta telah di sertifikasi oleh Balai Pengawasan dan Sertifikasi Benih (BPSB). Sedangkan benih yang tidak unggul merupakan benih dari varietas lokal atau dari hasil penangkaran sendiri yang telah dipilih dan dianggap memenuhi syarat untuk dijadikan benih padi oleh petani tanpa melalui proses pengawasan dan sertifikasi dari BPSB. Pengguna bibit padi unggul mendatangkan banyak keuntungan diantaranya meningkatkan mutu hasil yang nantinya berpengaruh terhadap peningkatan pendapat petani.

Untuk membantu dalam proses pencarian maka digunakan metode K-Means yang dipergunakan untuk mengelompokan data nonhierarki (sekatan) yang mempartisi data yang ada kedalam dua bentuk atau kelompok. Ada pun tujuan pengelompokan data ini adalah untuk meminimalkan fungsi objektif yang di set dalam proses pengelompokan yang pada umumnya berusaha meminimalkan variasi didalam suatu kelompok

Copyright © 2018 STMIK Triguna Dharma.
All rights reserved.

First Author

Nama : Suardi Yakub, SE, S.Kom, MM

Program Studi : Sistem Informasi STMIK Triguna Dharma

E-Mail :

1. PENDAHULUAN

Indonesia merupakan Negara dengan jumlah penduduk terpadat, setiap tahunnya jumlah penduduk di Indonesia terus meningkat dan peningkatan jumlah penduduk diikuti oleh peningkatan konsumsi beras nasional. Penduduk Indonesia mengkonsumsi beras sebagai makanan pokok. Produksi pangan harus terus ditingkatkan bagi pemenuhan kebutuhan pangan nasional. Peningkatan produksi tanaman pangan khususnya padi harus sejalan dengan laju pertumbuhan penduduk Indonesia agar pemenuhan kebutuhan pangan tercapai.

Perbedaan antara benih unggul dengan benih padi tidak unggul terletak pada proses sertifikasi, dimana benih bibit unggul di proses dan di pelihara sedemikian rupa sehingga tingkat kemurnian varietas dapat terpelihara dan memenuhi standar mutu benih yang ditetapkan serta telah di sertifikasi oleh Balai Pengawasan dan Sertifikasi Benih (BPSB). Sedangkan benih yang tidak unggul merupakan benih dari varietas lokal atau dari hasil penangkaran sendiri yang telah dipilih dan dianggap memenuhi syarat untuk dijadikan benih padi oleh petani tanpa melalui proses pengawasan dan sertifikasi dari BPSB. Pengguna bibit padi unggul mendatangkan banyak keuntungan diantaranya meningkatkan mutu hasil yang nantinya berpengaruh terhadap peningkatan pendapat petani. PT. Sang Hyang Seri (Persero) merupakan salah satu produsen benih padi yang telah berkembang di Indonesia dan merupakan penyumbang benih bagi pemenuhan

Kebutuhan benih unggul nasional. PT. SHS didirikan oleh pemerintah pada tahun 1971 dengan status semi-swasta sebagai Badan Usaha Milik Negara (BUMN) untuk mendampingi balai-balai benih dalam memproduksi benih. Salah satu lokasi penangkaran benih padi PT. Sang Hyang Seri terletak di Tanjung Morawa, Kabupaten Deli Serdang, Sumatera Utara. Data mining adalah serangkaian proses untuk mengenali nilai tambah berupa informasi yang selama ini tidak diketahui secara manual dari suatu basis data dengan

menggunakan penggalian pola-pola dari data dengan tujuan untuk memanipulasi data menjadi informasi yang lebih berharga yang diperoleh dengan cara mengekstraksi dan mengenali pola yang penting atau menarik data yang terdapat dalam basis data. Untuk membantu dalam proses pencarian maka digunakan metode K-Means yang dipergunakan untuk mengelompokkan data nonhierarki (sekatan) yang mempartisi data yang ada kedalam dua bentuk atau kelompok. Ada pun tujuan pengelompokan data ini adalah untuk meminimalkan fungsi objektif yang di set dalam proses pengelompokan yang pada umumnya berusaha meminimalkan variasi didalam suatu kelompok

2. LANDASAN TEORITIS

2.1 Bibit Padi

Benih merupakan salah satu komponen utama yang berperan penting dalam peningkatan kualitas produk padi, karenanya penggunaan benih varietas unggul yang bermutu (berlabel) sangat dianjurkan. Hal ini terkait dengan sifat-sifat yang dimiliki oleh varietas unggul, antara lain: berdaya hasil tinggi, tahan terhadap hama penyakit, dan rasa nasi enak (pulen). Faktor yang mempengaruhi produksi benih padi yaitu dengan adanya ketersediaan benih unggul bermutu tinggi bagi petani dalam melakukan kegiatan usaha tani merupakan syarat penting dalam peningkatan hasil dan kualitas produksi. Benih yang memiliki mutu baik sangatlah diperlukan oleh petani maupun penangkaran benih. Agar petani maupun penangkaran benih tidak merasa dirugikan serta mereka memiliki jaminan kualitas atas benih yang digunakannya, maka diajarkan menggunakan benih unggul sangatlah penting bagi benih unggul sudah ditetapkan kelas-kelas benih unggul.

Menurut Sadjad (1993) menyatakan bahwa Benih bermutu harus memenuhi criteria 7 tepat yaitu tepat varietas, tepat mutu, tepat jumlah, tepat waktu, tepat tempat, tepat harga dan tepat pelayanan. Hasil benih ini diberi sertifikat, sehingga dinamakan benih bersertifikat. Benih bersertifikat adalah benih yang proses produksinya menerapkan cara dan persyaratan tertentu sesuai dengan ketentuan standar mutu benih Balai Pengawasan dan Sertifikasi Benih (BPSB) melalui label benih bersertifikat..

2.2 Data Mining

Data Mining adalah bidang yang sepenuhnya menggunakan apa yang dihasilkan oleh data warehouse, bersama dengan bidang yang menangani masalah pelaporan dan manajemen data. Sementara, data warehouse sendiri bertugas untuk menarik /meng-query data dari basis data mentah untuk memberikan hasil data yang nantinya digunakan oleh bidang yang menangani manajemen, pelaporan, dan data mining. Dengan data mining inilah, penggalian informasi baru dapat dilakukan dengan bekal data mentah yang diberikan oleh data warehouse.

Hasil yang diberikan oleh ketiga bidang tersebut berguna untuk mendukung aktivitas bisnis cerdas (business intelligence). Pekerjaan yang berkaitan dengan data mining dapat dibagi menjadi empat kelompok, yaitu model prediksi (prediction modelling), analisis kelompok (cluster analysis), analisis asosiasi (association analysis), dan deteksi anomaly (anomaly detection).

2.3 K-Means

K-Mean merupakan salah satu metode pengelompokan data nonhierarki (sekatan) yang berusaha mempartisi data yang ada kedalam bentuk dua atau lebih kelompok. Metode ini mempartisi data ke dalam kelompok sehingga data berkarakteristik sama di masukan ke dalam satu kelompok yang sama dan data yang berkarakteristik berbeda dikelompokkan kedalam kelompok yang lain. Adapun tujuan pengelompokan data ini adalah untuk meminimalkan fungsi objektif yang diset dalam proses pengelompokan, yang pada umumnya berusaha meminimalkan variasi didalam suatu kelompok dan memaksimalkan variasi antar kelompok.

Lokasi sentroid (titik pusat) setiap kelompok yang diambil dari rata-rata (mean) semua nilai data pada setiap fitur nya harus dihitung kembali. Jika M menyatakan jumlah data dalam sebuah kelompok, i menyatakan fitur ke- i dalam sebuah kelompok, dan p menyatakan dimensi data, untuk menghitung sentroid fitur ke- i digunakan formula :

$$C_i = \frac{1}{M} \sum_{j=1}^M x_j$$

Formula tersebut dilakukan sebanyak p dimensi sehingga i mulai dari 1 sampai p . Menurut eko prasetyo, 2012:179 ada beberapa cara yang dapat digunakan untuk mengukur jarak data kepusat kelompok, diantaranya Euclidean (Bezdek,1981) Manhattan/city block (miyamoto dan agusta,1995), dan Minkowsky (Miyamoto dan Agusta,1995). Masing-masing cara mempunyai kelebihan dan kekurangan. Pengukuran jarak pada ruang jarak (*distance space*) Euclidean menggunakan formula:

$$D(X_2, X_1) = \|X_2 - X_1\|_2 = \sqrt{\sum_{j=1}^p |x_{2j} - x_{1j}|^2}$$

D adalah jarak antar data x_2 dan x_1 , dan $| \cdot |$ adalah nilai mutlak. Pengukuran jarak pada ruang jarak Manhattan menggunakan formula.

$$D(X_2, X_1) = \|X_2 - X_1\|_2 = \sum_{j=1}^p |x_{2j} - x_{1j}|$$

Pengukuran jarak pada ruang jarak Minkowsky menggunakan formula.

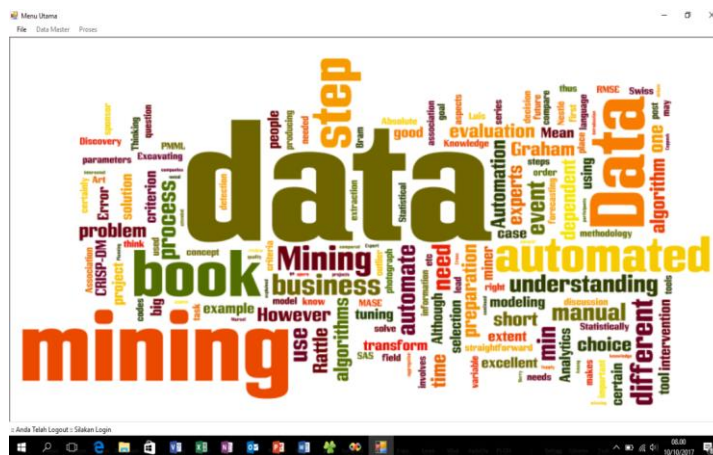
$$D(X_2, X_1) = \|X_2 - X_1\|_2 = \sqrt{\sum_{j=1}^p |x_{2j} - x_{1j}|^\lambda}$$

λ adalah parameter jarak Minkowsky. Secara umum, λ merupakan parameter penentu dalam karakteristik jarak. Jika $\lambda = 1$ ruang jarak pada Minkowsky sama dengan Manhattan. Jika $\lambda = 2$, ruang jaraknya akan sama dengan Euclidean, jika $\lambda = \infty$, ruang jaraknya akan sama dengan ruang jarak Chebyshev. Namun demikian, cara yang paling banyak digunakan Euclidean dan Manhattan. Euclidean menjadi pilihan jika kita ingin memberikan jarak terpendek antara dua titik (jarak lurus), seperti yang ditunjukkan, sedangkan Manhattan memberikan jarak terjauh pada dua data. Manhattan juga sering digunakan karena kemampuannya dalam mendeteksi keadaan khusus, seperti keberadaan outlier, dengan lebih baik.

3. ANALISIS DAN HASIL

Untuk menguji kebenaran dari hasil pengolahan data yang dikerjakan secara manual pada Bab III tersebut digunakanlah *software Microsoft Visual Studio 2008*. Berikut ini adalah tahapan dari pengaplikasian *Data Mining* dengan menggunakan metode *K-Means* dalam menentukan persediaan stok bibit padi unggul pada PT Sang Hyang Seri Medan :

1. Pertama kali menjalankan aplikasi maka akan tampil menu utama dimana menu pada menu utama masih belum dapat digunakan karena pengguna belum melakukan *login* kesistem. Dalam *form* menu utama terdapat dua tampilan yaitu menu *Home* dan Aplikasi, menu *home* digunakan untuk menampung menu *login*, keluar program, tentang sistem, dan tentang *programmer*. Berikut ini adalah tampilan *form* utama beranda :



Gambar 1 Tampilan *Interface Form* MenuUtama

2. *Interface Form Login*

Tampilan *form login* dapat dipanggil ketika pengguna memilih menu *home* dan terdapat *button login*. Berikut ini adalah gambar tampilan *form login*;

Gambar 2 Tampilan *Interface Form Login*

Dari gambar tersebut diatas terdapat dua *textbox* yang berfungsi sebagai tempat untuk melakukan inputan *user name* serta *password* dan dua *button* yang berfungsi untuk menjadi tombol *login* dan tombol batal *login*.

3. *Form* masukan *interface* sistem untuk Input Data Penjualan sebagai berikut :

ID	Tanggal Penjualan	Pandan Wangi	Bestari	Caredek Merah
1	01/02/2015	2167	4300	1237
2	02/02/2015	2402	8342	4340
3	03/02/2015	827	1278	2467
4	04/02/2015	3043	4320	3043
5	05/02/2015	2876	3567	3107
6	06/02/2015	578	0	1789
7	07/02/2015	365	0	598
8	08/02/2015	1435	2907	3035
9	09/02/2015	2985	3864	3278
10	10/02/2015	207	406	382
11	11/02/2015	158	0	739
12	12/02/2015	1639	2692	2038
13	13/02/2015	710	0	1209
14	14/02/2015	4231	0	6321
15	15/02/2015	4389	6321	5321
16	16/02/2015	12587	17939	20932
17	17/02/2015	5302	7321	6456
18	18/02/2015	2031	3201	4092
19	19/02/2015	372	459	621
20	20/02/2015	271	329	421

Gambar 3 Tampilan *Interface Form Input* Data Penjualan

Dari gambar diatas terdapat lima buah *textbox* yang berfungsi sebagai tempat menampung data yang nantinya akan di *input* kedalam *database* kemudian terdapat lima *button* yang digunakan sebagai perintah untuk melakukan pengolahan data seperti simpan, hapus, ubah, dan tutup *form*.

4. Selanjutnya tampilan *interface form* hasil perhitungan *K-Means* sebagai berikut :

Gambar 4 Tampilan *Interface Form Hasil Perhitungan Metode K-Means*

Dari gambar diatas terdapat Sembilan *textbox* yang nantinya digunakan untuk menampung nilai *cluster* kemudian terdapat tombol sebanyak tiga buah yang masing-masing digunakan untuk melakukan proses perhitungan metode *K-Means* menutup program dan melakukan pencetakan laporan perhitungan *K-Means*.

Setelah melakukan proses implementasi, proses selanjutnya adalah uji coba dengan tujuan untuk mengetahui bahwa hasil perancangan di bab III sesuai dengan hasil yang ditampilkan pada aplikasi. Dalam melakukan pengujian maka dibutuhkan data untuk menjadi bahan yang digunakan dalam perhitungan algoritma *K-Means*. Adapun data yang dibutuhkan dalam perhitungan algoritma *K-Means* yaitu :

1. Tampilan Form Input Data Penjualan

Form Input

Input Data Penjualan Bibit Padi

ID Bibit Padi:

Tanggal Penjualan:

Jumlah Penjualan Bibit Padi:

Pandan Wangi:

Bestari:

Caredek Merah:

Tambah Simpan Hapus Batal Tutup

Data Penjualan Bibit Padi

ID	Tanggal Penjualan	Pandan Wangi	Bestari	Caredek Merah
1	01/02/2015	2167	4300	1237
2	02/02/2015	2402	8342	4340
3	03/02/2015	827	1278	2467
4	04/02/2015	3043	4320	3043
5	05/02/2015	2876	3567	3107
6	06/02/2015	578	0	1789
7	07/02/2015	365	0	598
8	08/02/2015	1435	2907	3035
9	09/02/2015	2985	3864	3278
10	10/02/2015	207	406	382
11	11/02/2015	158	0	739
12	12/02/2015	1639	2692	2038
13	13/02/2015	710	0	1209
14	14/02/2015	4231	0	6321
15	15/02/2015	4389	6321	5321
16	16/02/2015	12587	17939	20932
17	17/02/2015	5302	7321	6456
18	18/02/2015	2031	3201	4092
19	19/02/2015	372	459	621
20	20/02/2015	271	329	421

Gambar 5 Tampilan *Interface Form* Input Data Penjualan

2. Tampilan Form Proses Data Metode *K-Means*.

Proses KMEANS

Menentukan Pusat Awal Cluster

Pandan Wangi: 12587
Bestari: 17939
Caredek Merah: 20932

Pandan Wangi: 5302
Bestari: 7321
Caredek Merah: 6456

Pandan Wangi: 37
Bestari: 92
Caredek Merah: 105

Iterasi ke :
Proses Lanjut
Tutup
Laporan

Perhitungan Jarak Antar Cluster

Tanggal Penjualan	Pandan Wangi	Bestari	Caredek Merah	C1	C2	C3
01/02/2015	2167	4300	1237	26124.54	6796.52	4850
02/02/2015	2402	8342	4340	21705.55	3732.28	9570
03/02/2015	827	1278	2467	27510.78	8512.09	2758
04/02/2015	3043	4320	3043	24425.02	5075.2	5961
05/02/2015	2876	3567	3107	24871.44	5585.14	5388
06/02/2015	578	0	1789	28852.7	9884.03	1771
07/02/2015	365	0	598	29743.18	10596.59	599.2
08/02/2015	1435	2907	3035	25896.54	6792.67	4296
09/02/2015	2985	3864	3278	24535.03	5236.32	5743
10/02/2015	207	406	382	29714.86	10519.97	451.5
11/02/2015	158	0	739	29732.89	10618	651.5
12/02/2015	1639	2692	2038	26532.93	7373.19	3614
13/02/2015	710	0	1209	29186.77	10110.12	1296
14/02/2015	4231	0	6321	24599.02	7400.16	7499
15/02/2015	4389	6321	5321	21116.07	1766.86	9216
16/02/2015	12587	17939	20932	0	19374.41	3016
17/02/2015	5302	7321	6456	19374.41	0	1096
18/02/2015	2031	3201	4092	24743.15	5767.35	5434
19/02/2015	372	459	621	29449.84	10288.36	716.3
20/02/2015	271	329	421	29706.85	10517.62	456.1
21/02/2015	432	921	823	29012.55	9818.74	1165
22/02/2015	37	92	105	30162.59	10968.77	0
23/02/2015	2051	4731	5710	23388.74	4655.06	6873

Pengeklompokan Data

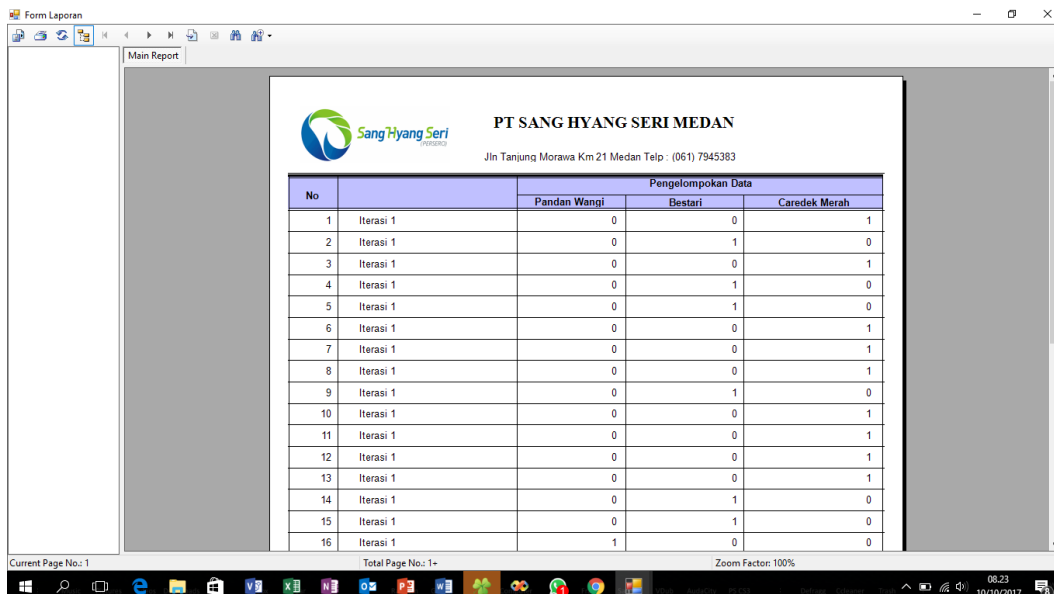
	C1	C2	C3
1			1
2			1
3			1
4			1
5			1
6			1
7			1
8			1
9			1
10			1
11			1
12			1
13			1
14			1
15			1
16			1
17			1
18			1
19			1
20			1
21			1
22			1
23			1

© Login Sebagai : admin

Gambar 6 Tampilan *Interface Form* Proses

3. Tampilan Form Laporan

Setelah melakukan pengujian, untuk hasil/output berupa nilai hasil perhitungan dan keputusan yang diberikan sistem, bentuk laporan hasil perhitungan algoritma *K-Means* adalah sebagai berikut



Form Laporan

Main Report

PT SANG HYANG SERI MEDAN
 Jln Tanjung Morawa Km 21 Medan Telp : (061) 7945383

No		Pengelompokan Data		
		Pandan Wangi	Bestari	Caredek Merah
1	Iterasi 1	0	0	1
2	Iterasi 1	0	1	0
3	Iterasi 1	0	0	1
4	Iterasi 1	0	1	0
5	Iterasi 1	0	1	0
6	Iterasi 1	0	0	1
7	Iterasi 1	0	0	1
8	Iterasi 1	0	0	1
9	Iterasi 1	0	1	0
10	Iterasi 1	0	0	1
11	Iterasi 1	0	0	1
12	Iterasi 1	0	0	1
13	Iterasi 1	0	0	1
14	Iterasi 1	0	1	0
15	Iterasi 1	0	1	0
16	Iterasi 1	1	0	0

Current Page No: 1 Total Page No: 1- Zoom Factor: 100%

Gambar 7 Tampilan Interface Form Laporan Perhitungan Metode K-Means

4. KESIMPULAN

Dari hasil pembahasan pada bab sebelumnya dan pengamatan yang telah dilakukan maka dapat diambil kesimpulan diantaranya adalah sebagai berikut:

- 1 Sistem yang telah dibangun ini dapat menganalisis data penjualan Bibit Unggul dengan menerapkan metode K-Means.
- 2 Dapat menerapkan hasil perhitungan dengan metode K-Means dalam pengelolaan data persediaan Bibit Padi Pada PT.Sang Hyang Seri.
- 3 Dapat mengimplementasikan Metode K-Means dalam sebuah sistem yang nantinya akan digunakan untuk menganalisis data penjualan Bibit Padi.

DAFTAR PUSTAKA

- Prasetyo,E.2012. Data Mining Konsep dan Aplikasi Menggunakan Matlab,Yogyakarta : Andi Offset.
 Sadjad, Syamsul. 1993. Dari Benih Kepada Benih. PT.Gramedia Widiasarana Indonesia. Jakarta.
 Rosa,A.S., & M. 2013. Rekayasa Perangkat Lunak Terstruktur dan Berorientas Objek. Bandung : Informatika
 Sulindawaty, Fathoni, M (Agustus 2010). Pengantar analisa Perancangan Sistem Saintikom, 9(2), 491-508
 Hendrayudi (Maret 2011) Dasr-dasar Pemrograman Microsoft Visual Basic 2008. Bandung

BIOGRAFI PENULIS



Suardi Yakub, SE, S.Kom, MM merupakan Dosen Tetap STMIK Triguna Dharma yang aktif mengampu mata kuliah Prinsip Manajemen Bisnis, Techno Enterpreuner, Metode Penelitian, Teknik Pemasaram Statistika, Akuntansi dan Komputer Akuntansi. Pria kelahiran Pariaman 06 April 1966 ini saat ini menduduki jabatan Wakil Ketua II Bidang Keuangan. Tamat S1 dan S2 di Universitas Islam Sumatera Utara bidang manajemen, dan beliau mengambil pendidikan kembali pada bidang ilmu komputer di Sekolah Tinggi Manajemen Informatika dan Komputer ITMI Medan