

**ANALISIS *CLUSTER* DENGAN ALGORITMA
K-MEANS DAN *FUZZY C-MEANS CLUSTERING*
UNTUK PENGELOMPOKAN DATA OBLIGASI
KORPORASI**



=====
SKRIPSI
=====

Disusun Oleh :

DESY RAHMAWATI NINGRAT

24010212120009

**DEPARTEMEN STATISTIKA
FAKULTAS SAINS DAN MATEMATIKA
UNIVERSITAS DIPONEGORO
SEMARANG**

2016

**ANALISIS *CLUSTER* DENGAN ALGORITMA *K-MEANS* DAN *FUZZY C-MEANS CLUSTERING* UNTUK PENGELOMPOKAN DATA OBLIGASI
KORPORASI**

Oleh:

DESY RAHMAWATI NINGRAT

24010212120009

Sebagai Salah Satu Syarat untuk Memperoleh Gelar
Sarjana Statistika pada Departemen Statistika

**DEPARTEMEN STATISTIKA
FAKULTAS SAINS DAN MATEMATIKA
UNIVERSITAS DIPONEGORO
SEMARANG**

2016

HALAMAN PENGESAHAN I

Judul : Analisis *Cluster* dengan Algoritma *K-Means* dan *Fuzzy C-Means Clustering* untuk Pengelompokan Data Obligasi Korporasi.

Nama : Desy Rahmawati Ningrat

NIM : 24010212120009

Telah diujikan pada sidang Tugas Akhir dan dinyatakan lulus pada tanggal 19 September 2016

Semarang, 28 September 2016

Mengetahui,

Ketua Departemen Statistika

ESM UNDIP

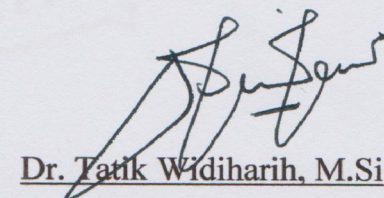


Dra. Hj. Dwi Ispriyanti, M.Si

NIP. 195709141986032001

Panitia Penguji Ujian Tugas Akhir

Ketua,



Dr. Tatik Widiharah, M.Si

NIP. 196109281986032002

HALAMAN PENGESAHAN II

Judul : Analisis *Cluster* dengan Algoritma *K-Means* dan *Fuzzy C-Means Clustering* untuk Pengelompokan Data Obligasi Korporasi.

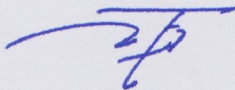
Nama : Desy Rahmawati Ningrat

NIM : 24010212120009

Telah diujikan pada sidang Tugas Akhir tanggal 19 September 2016.

Semarang, 28 September 2016

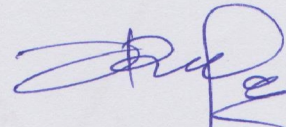
Dosen Pembimbing I



Di Asih I Maruddani, S.Si, M.Si

NIP. 197307111997022001

Dosen Pembimbing II



Triastuti Wuryandari, S.Si, M.Si

NIP. 197109061998032001

KATA PENGANTAR

Puji syukur penulis panjatkan kehadiran Allah SWT yang telah memberikan rahmat dan karunia-Nya sehingga Tugas Akhir ini dapat terselesaikan. Tugas Akhir yang berjudul “Analisis *Cluster* dengan Algoritma *K-Means* dan *Fuzzy C-Means Clustering* untuk Pengelompokan Data Obligasi Korporasi” ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana Statistika pada Departemen Statistika Fakultas Sains dan Matematika Universitas Diponegoro.

Banyak pihak yang telah membantu dalam penyelesaian Tugas Akhir ini. Oleh karena itu, rasa hormat dan terima kasih penulis ingin sampaikan kepada :

1. Ibu Dra. Dwi Ispriyanti, M.Si. selaku Ketua Departemen Statistika Fakultas Sains dan Matematika Universitas Diponegoro,
2. Ibu Di Asih I Maruddani, S.Si, M.Si. dan Ibu Triastuti Wuryandari, S.Si, M.Si. selaku dosen pembimbing I dan II yang telah berkenan atas bimbingan, arahan, dan waktunya yang diberikan kepada penulis hingga terselesaikannya tugas akhir ini.
3. Semua pihak yang tidak dapat penulis sebut satu persatu yang telah membantu dalam penyelesaian penulisan tugas akhir ini.

Penulis menyadari bahwa tugas akhir ini masih jauh dari sempurna. Oleh karena itu, kritik dan saran yang membangun sangat penulis harapkan demi perbaikan dalam kesempatan berikutnya.

Semarang, September 2016

Penulis

ABSTRAK

Analisis *cluster* adalah metode pengelompokan data (objek) yang didasarkan hanya pada informasi yang ditemukan dalam data yang menggambarkan objek tersebut dan hubungan diantaranya. Analisis *cluster* bertujuan agar objek-objek yang tergabung dalam sebuah *cluster* merupakan objek-objek yang mirip (atau berhubungan) satu sama lain dan berbeda (tidak berhubungan) dengan objek dalam *cluster* yang lain. Pada penelitian ini digunakan dua metode pengelompokan yakni *Fuzzy C-Means* (FCM) dan *K-Means clustering*. Data yang digunakan adalah data obligasi 357 perusahaan pada tanggal 1 Desember 2015. Variabel-variabel yang digunakan terdiri dari *coupon rate* (tingkat kupon), *time to maturity* (waktu sampai dengan jatuh tempo), *yield* (tingkat pengembalian), dan *rating* (peringkat) masing-masing perusahaan.. Penentuan banyaknya *cluster* optimum dilakukan dengan perhitungan indeks validitas Xie Beni pada metode FCM. Setelah diperoleh banyaknya *cluster* optimum, dilakukan tahap evaluasi yakni membandingkan metode FCM dan metode *K-Means* dengan memperhatikan nilai rasio rata-rata simpangan baku di dalam *cluster* dan rata-rata simpangan baku antar *cluster* (Sw/Sb) masing-masing metode. Metode dengan nilai rasio Sw/Sb terkecil yang akan dipilih sebagai metode terbaik. Berdasarkan indeks validitas Xie Beni, disimpulkan jumlah *cluster* optimum adalah 10 *cluster*. Hasil pengelompokan yang digunakan adalah *K-Means clustering* karena memiliki nilai rasio Sw/Sb terkecil dibandingkan FCM yakni 0,6651. Kemudian hasil *K-Means clustering* dianalisis untuk mengetahui interpretasi dan karakteristik dari masing-masing *cluster* yang terbentuk.

Kata kunci: Analisis *cluster*, *coupon rate*, *time to maturity*, *yield*, *rating*, *Fuzzy C-Means*, *K-Means*, Indeks Xie Beni, Rasio Sw/Sb.

ABSTRACT

Cluster analysis is a method of grouping data (object) that are based on information that found in the data which describes the object and relation within. Cluster analysis aims to make the joined objects in the cluster are identical (or related) with one another and different (not related) to objects in another cluster. In this study used two method of grouping; Fuzzy C-Means and K-Means Clustering. The data used in this research had been using 357 corporate bonds data on December 1st, 2015. The variables used in this study consist of coupon rate, time to maturity, yield and rating of each corporate. The determination of the number of optimum clusters performed by Xie Beni index of validity calculation at FCM method. Having obtained the optimum number of clusters, evaluation step was conducted by comparing FCM method to K-Means method with noticing the average of standard deviation in the clusters and the average of standard deviation inter-clusters (Sw/Sb) from each method. Method with the smallest Sw/Sb ratio value would get chosen as the best method. Based on the validity index Xie Beni, the most optimum number of cluster is 10 because it has the smallest Sw/Sb ratio value compared to FCM, the value is 0,6651. Afterwards, the result of K-Means clustering is analyzed to determined the interpretation and characteristics of each formed clusters.

Keyword: Cluster Analysis, coupon rate, time to maturity, yield, rating, Fuzzy C-Means, K-Means, Xie Beni Index, Sw/Sb ratio.

DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
HALAMAN PENGESAHAN I	ii
HALAMAN PENGESAHAN II	iii
KATA PENGANTAR	iv
ABSTRAK	v
ABSTRACT	vi
DAFTAR ISI	vii
DAFTAR TABEL.....	x
DAFTAR LAMPIRAN.....	xi
BAB I PENDAHULUAN	
1.1 LatarBelakang	1
1.2 RumusanMasalah	5
1.3 BatasanMasalah.....	5
1.4 Tujuan Penelitian	5
BAB II TINJAUAN PUSTAKA	
2.1 Obligasi Perusahaan	7
2.2 Karakteristik Obligasi	8
2.2.1 Nilai Penerbitan Obligasi (Jumlah Pinjaman dana)	8
2.2.2 Jangka Waktu Obligasi	9

2.2.3	Tingkat Suku Bunga	9
2.2.4	Jadwal Pembayaran Suku Bunga	9
2.2.5	Jaminan	10
2.3	Jenis-Jenis Obligasi	10
2.3.1	Berdasarkan Penerbitnya (emiten)	10
2.3.2	Berdasarkan Suku Bunga Obligasi	12
2.4	Struktur Obligasi	13
2.5	Keunggulan Obligasi	15
2.6	<i>Rating</i> Obligasi.....	16
2.7	Analisis <i>Cluster</i>	19
2.7.1	Pemilihan Ukuran Jarak	20
2.7.2	Asumsi Analisis <i>Cluster</i>	21
2.8	Memilih Prosedur Analisis <i>Cluster</i>	24
2.9	Pengelompokan dengan FCM dan <i>K-Means</i>	25
2.10	Penentuan Jumlah <i>Cluster</i> Optimum (Indeks Xie Beni)	30
2.11	Tahap Evaluasi Penentuan Metode Terbaik	31
BAB III METODOLOGI PENELITIAN		
3.1	Jenis dan Sumber Data	32
3.2	Variabel Data	32
3.3	Langkah Analisis	33
3.4	Diagram Alir Pengolahan Data	35
BAB IV HASIL DAN PEMBAHASAN		
4.1	Analisis Deskriptif.....	36

4.2	Pengelompokan dengan Algoritma FCM.....	36
4.3	Pengelompokan dengan Algoritma <i>K-Means</i>	51
4.3.1	Pengujian Asumsi	51
4.3.2	Hasil Pengelompokan <i>K-Means</i>	52
4.4	Penentuan Metode Terbaik.....	53
4.5	Interpretasi Hasil Pengelompokan.....	54
4.6	Karakteristik <i>Cluster</i>	68
BAB VPENUTUP		
5.1	Kesimpulan.....	75
5.2	Saran.....	76
DAFTAR PUSTAKA		77
LAMPIRAN.....		79

DAFTAR TABEL

	Halaman
Tabel 1 <i>Rating</i> Obligasi di Indonesia (PEFINDO)	18
Tabel 2 Skoring <i>Rating</i> Data Penelitian	19
Tabel 3 Statistik Deskriptif Obligasi Korporasi Tanggal 1 Desember 2015	36
Tabel 4 Rasio Sw/Sb Algoritma FCM	49
Tabel 5 Hasil Pengelompokan FCM	50
Tabel 6 VIF Masing-masing Variabel.....	52
Tabel 7 Rasio Sw/Sb Algoritma <i>K-Means</i>	53
Tabel 8 Rasio Sw/Sb Metode Terbaik	53

DAFTAR LAMPIRAN

	Halaman
Lampiran1 Data Obligasi Korporasi Tanggal 1 Desember 2015	79
Lampiran2 Syntax Program Matlab	101
Lampiran3 Output Matlab Metode FCM dan <i>K-Means</i>	107
Lampiran4 Output Pengujian Asumsi Multikolinearitas.....	116
Lampiran5 Rata-rata Variabel Penelitian	117

BAB I

PENDAHULUAN

1.1 Latar Belakang

Data merupakan salah satu sumber yang dapat digunakan untuk mendapatkan suatu informasi. Akan tetapi, tidak jarang kumpulan data tersebut dibiarkan begitu saja, sehingga diperlukan suatu metode yang dapat dipakai untuk menggali informasi sebanyak mungkin dari data tersebut. *Data Mining* sebagai salah satu ilmu di bidang statistika, dapat digunakan untuk mengekstraksi informasi berharga yang sebelumnya tidak diketahui dari suatu *database*. Sebagai contoh adalah data obligasi korporasi yang akan digunakan pada Tugas Akhir ini.

Obligasi adalah sekuritas yang memuat janji untuk memberikan pembayaran tetap menurut jadwal yang telah ditetapkan (Rahardjo, 2003). Informasi tentang perkembangan harga obligasi dapat dijumpai pada laporan *Indonesia Bond Pricing Agency* (IBPA) yang dapat diakses melalui internet (www.ibpa.co.id). Melalui informasi tersebut dapat diketahui pergerakan harga obligasi dan estimasi pergerakan tingkat bunga yang sangat berpengaruh terhadap harga obligasi. Obligasi sebagai instrumen investasi diperkirakan akan terus bertumbuh dari segi jumlah *issuer* maupun volume perdagangannya. Oleh karena itu, informasi yang berkaitan dengan obligasi harus diketahui oleh masyarakat secara luas. Selain informasi yang dikeluarkan oleh bursa dan agen-agen investasi lain, investor juga perlu melengkapi diri dengan kemampuan

analisis dan teknik-teknik perhitungan yang berkaitan dengan investasi obligasi, dengan demikian investor memahami tentang risiko dan tingkat pengembalian (*return*) yang terkandung dalam investasi obligasi tersebut (Zubir, 2012).

Menurut Zubir (2012), risiko yang terkandung dalam sebuah obligasi meliputi risiko tingkat bunga (*interest rate risk*) dan risiko gagal bayar (*default risk*). Walaupun obligasi dikelompokkan sebagai sekuritas berpendapatan tetap (*fixed income securities*), tetapi kemungkinan hasil yang diharapkan tidak sesuai dengan kenyataannya tetap ada. Meskipun obligasi membayarkan kupon yang tetap secara periodik, namun nilai kupon tersebut dapat tidak sesuai dengan yang dijanjikan karena perusahaan yang menerbitkan obligasi tersebut (*issuer*) mengalami masalah keuangan. Selain itu, nilai kupon bila diinvestasikan kembali akan menurun bila tingkat bunga juga turun dan meningkat bila tingkat bunga naik. Sedangkan harga obligasi akan naik bila tingkat bunga turun dan menurun bila tingkat bunga naik.

Besarnya risiko gagal bayar obligasi yang dikeluarkan oleh perusahaan (*corporate bond*) sangat bergantung pada kondisi perusahaan yang bersangkutan. Untuk menjaga keamanan obligasi dan melindungi investor, obligasi perusahaan diberi peringkat (*rating*) oleh lembaga atau *agency* yang independen. Lembaga pemeringkat dari Amerika seperti Standar & Poor dan Moody's, dan Indonesia seperti Pefindo dan Casnic. Obligasi perusahaan yang berperingkat tinggi, misalnya AAA menunjukkan bahwa gagal bayar obligasi tersebut kecil. Sebaliknya obligasi perusahaan dengan peringkat rendah, seperti B mempunyai probabilitas gagal bayar yang besar. Untuk mengkompensasi risiko gagal bayar

tersebut, obligasi dengan peringkat rendah akan dijual murah atau investor meminta *yield to maturity* yang besar. Walaupun obligasi perusahaan sudah dinyatakan peringkatnya dan investor juga sudah mengetahuinya, pilihan untuk membeli obligasi dengan peringkat tinggi atau rendah ada di tangan investor, dalam hal ini berlaku "*high risk, high return*".

Berdasarkan informasi yang dapat digali dari data obligasi korporasi laporan *Indonesia Bond Pricing Agency* (IBPA) tersebut maka dapat dilakukan proses pengelompokan obligasi. Hal ini dilakukan untuk mendukung strategi pemilihan investasi obligasi yang bisa jadi berbeda untuk tiap kelompoknya. Data obligasi ini terdiri dari beberapa atribut dengan jumlah *record* yang banyak sehingga diperlukan suatu proses yang dapat mengelompokkan data tersebut, yaitu *clustering*. Penggunaan *clustering* diharapkan dapat memberikan prediksi pengelompokan obligasi tersebut.

Salah satu metode *clustering* yang dapat digunakan untuk mengelompokkan data adalah *Fuzzy C-Means* (FCM) dan *K-Means clustering*. FCM dan *K-Means clustering* merupakan metode pengelompokan data nonhierarki (sekatan). FCM adalah suatu teknik pengelompokan data dimana keberadaan tiap titik data dalam suatu kelompok (*cluster*) ditentukan oleh derajat keanggotaan. Teknik ini pertama kali diperkenalkan oleh Jim Bezdek pada tahun 1981. Konsep dasar FCM yaitu pertama menentukan pusat *cluster* yang akan menandai lokasi rata-rata untuk tiap-tiap *cluster*. Pada kondisi awal pusat *cluster* masih belum akurat. Tiap-tiap data pada setiap *cluster* memiliki derajat keanggotaan. Dengan cara memperbaiki pusat *cluster* dan nilai keanggotaan tiap-tiap data secara berulang, maka akan didapat

bahwa pusat *cluster* menuju lokasi yang tepat. Konsep perulangan ini sama seperti metode *K-Means* yakni didasarkan pada minimalisasi fungsi objektif (Kusumadewi, 2010). Dengan metode FCM ini, akan dianalisis banyaknya *cluster* optimum yang akan mengelompokkan data obligasi korporasi ke dalam *cluster-cluster* yang berbeda menggunakan Indeks Xie Beni.

K-Means Clustering merupakan metode yang berusaha mempartisi data yang ada ke dalam bentuk dua atau lebih kelompok (Prasetyo, 2012). Metode ini mempartisi data ke dalam kelompok (*cluster*) sehingga data berkarakteristik sama dimasukkan ke dalam satu *cluster* yang sama dan data yang berbeda dikelompokkan ke dalam *cluster* yang lain. Tujuan pengelompokan dengan *K-Means Clustering* adalah untuk meminimalkan fungsi objektif yang dilakukan dalam proses pengelompokan, dengan tujuan meminimalkan variasi di dalam suatu *cluster* dan memaksimalkan variasi antar *cluster*. Pada penelitian ini setelah menganalisis dengan metode FCM dan diperoleh banyak *cluster* optimum berdasarkan indeks Xie Beni, kemudian melakukan tahap evaluasi yaitu dengan membandingkan hasil pengelompokan dengan metode *K-Means*. Tahap penentuan metode terbaik dilakukan dengan membandingkan rasio simpangan baku didalam *cluster* (S_w) terhadap rasio simpangan baku antar *cluster* (S_b) pada masing-masing metode. Nilai rasio S_w/S_b yang terkecil yang akan dipilih sebagai metode terbaik.

1.1 Rumusan Masalah

Berdasarkan latar belakang masalah, maka permasalahan pada Tugas Akhir ini, adalah :

1. Berapa jumlah *cluster* optimum yang terbentuk berdasarkan perhitungan indeks validitas Xie Beni?
2. Bagaimana perbandingan hasil pengelompokan dengan FCM dan *K-Means Clustering* untuk mengelompokkan data obligasi korporasi?
3. Bagaimana interpretasi terhadap *cluster* yang terbentuk dari data obligasi korporasi berdasarkan *coupon rate*, *time to maturity*, *yield*, dan *rating* dari masing-masing perusahaan berdasarkan metode terbaik?

1.3 Batasan Masalah

Pada Tugas Akhir ini, batasan masalah yang digunakan adalah :

Data obligasi korporasi yang diperoleh dari laporan *Indonesia Bond Pricing Agency* (IBPA) pada tanggal 1 Desember 2015.

1.4 Tujuan Penelitian

Secara umum tujuan yang ingin dicapai dalam Tugas Akhir ini adalah :

1. Mengetahui jumlah *cluster* optimum berdasarkan perhitungan indeks validitas Xie Beni.

2. Membandingkan antara metode FCM dan *K-Means Clustering*, untuk kemudian dipilih metode mana yang terbaik berdasarkan nilai rasio S_w/S_b terkecil.
3. Menginterpretasi hasil pengelompokan obligasi berdasarkan *coupon rate*, *time to maturity*, *yield*, dan *rating* masing-masing perusahaan menggunakan metode terbaik.