

Implementasi *K-Means Clustering* untuk Pengelompokan Analisis Rasio Profitabilitas dalam *Working Capital*

N. Harianto Kristanto.¹, Andreas Christopher L.A.², Halim Budi S.³

Abstrak— Rasio profitabilitas merupakan rasio untuk mengetahui kemampuan perusahaan dalam menghasilkan laba selama periode tertentu dan juga memberikan gambaran tentang tingkat efektifitas manajemen perusahaan dalam melaksanakan kegiatan operasinya. Rasio profitabilitas diperoleh dari laporan keuangan tahunan perusahaan dengan melakukan perhitungan dari data penjualan yang terdapat di dalam laporan tahunan. Masalahnya bagaimana mengelompokkan emiten-emiten yang punya kemiripan dalam rasio profitabilitas..

Dalam penelitian ini akan dibangun sistem berbasis desktop dengan menerapkan metode *k-means clustering*. Sistem akan mengelompokkan rasio profitabilitas semua perusahaan pada sektor pertambangan dan sektor industri barang konsumsi dalam 1 tahun. Hasilnya berupa cluster dari perusahaan mana yang memiliki kesamaan paling mirip pada rasio profitabilitasnya, hasil cluster juga disajikan dalam bentuk diagram scatter. Pada penelitian ini dilakukan uji coba terhadap centroid 2 hingga centroid 5.

Setelah dilakukan uji coba terhadap jumlah centroidnya, dapat ditarik kesimpulan bahwa semakin banyak jumlah centroid dalam setiap proses clustering, maka makin spesifik kelompok cluster yang dihasilkan. Dengan demikian pengambilan kesimpulan kesamaan dalam kelompok cluster makin mudah.

Kata Kunci: rasio profitabilitas, *K-means Clustering*, diagram scatter

Abstract— Profitability ratio is the ratio to determine a company's ability to generate profits for a certain period and also give an idea of the effectiveness of the company's management in conducting operations. Profitability ratios derived from the annual financial statements of the company to perform the calculation of the sales data contained in the annual report. The problem is how to classify issuers that have a similarity in profitability.

In this research will build desktop-based system by applying the *k-means clustering* method. The system will group all the company's profitability ratios in the mining sector and the consumer goods industry sector within 1 year. The result is a cluster of companies which have in common the most similar to the ratio of profitability, cluster results are also presented in the form of a scatter diagram. In this study, carried out tests on the centroid 2 to 5.

After testing against centroidnya number, it can be deduced that the greater number of centroid in every process of clustering, the more specific the resulting cluster group. Thus the conclusions similarities in the cluster group more easily.

Keyword: profitability ratios, *K-means clustering*, scatter diagrams

I. PENDAHULUAN

Rasio profitabilitas merupakan rasio untuk mengetahui kemampuan perusahaan dalam menghasilkan laba selama periode tertentu dan juga memberikan gambaran tentang tingkat efektifitas manajemen perusahaan dalam melaksanakan kegiatan operasinya. Tingkat efektifitas manajemen disini dilihat dari laba yang dihasilkan terhadap penjualan dan investasi perusahaannya. Melalui rasio profitabilitas, pemilik perusahaan dapat mengetahui, apakah perusahaannya memperoleh keuntungan atau kerugian. Penggunaan rasio profitabilitas dapat dilakukan dengan menggunakan perbandingan antara berbagai komponen yang ada dilaporan keuangan neraca dan laporan laba rugi. Pengukuran dapat dilakukan untuk beberapa periode operasi. Tujuannya adalah agar terlihat perkembangan perusahaan dalam rentang waktu tertentu.

Penelitian ini dilakukan untuk membangun sistem berbasis desktop dengan menerapkan metode *k-means clustering*. Sistem akan mengelompokkan rasio profitabilitas semua perusahaan pada sektor pertambangan dan sektor industri barang konsumsi dalam 1 tahun. Hasilnya berupa cluster dari perusahaan mana yang memiliki kesamaan paling mirip pada rasio profitabilitasnya, hasil cluster juga disajikan dalam bentuk diagram scatter.

II. LANDASAN TEORI

A. Pasar Modal

Pasar modal merupakan pasar untuk berbagai instrumen keuangan jangka panjang yang bisa

1. Jurusan Sistem Informasi, Universitas Kristen Duta Wacana, Jl. Dr. Wahidin 5-25, Yogyakarta, Indonesia (telp 0274-563929; fax 0274-513235; email: harianto@staff.ukdw.ac.id)
2. Jurusan Sistem Informasi, Universitas Kristen Duta Wacana, Jl. Dr. Wahidin 5-25, Yogyakarta, Indonesia (telp 0274-563929; fax 0274-513235; email: pokecrot@gmail.com)
3. Jurusan Sistem Informasi, Universitas Kristen Duta Wacana, Jl. Dr. Wahidin 5-25, Yogyakarta, Indonesia (telp 0274-563929; fax. 0274-513235; email: hbudi@staff.ukdw.ac.id)

diperjualbelikan atau biasa disebut dengan sekuritas. Sekuritas itu sendiri dapat berupa surat utang (obligasi), saham, reksa dana, instrumen derivatif maupun instrumen lainnya. Secara sederhana, pasar modal adalah tempat bertemunya pembeli dan penjual efek. Pasar modal merupakan sarana pendanaan bagi perusahaan maupun institusi lain dan sebagai sarana bagi kegiatan berinvestasi. Pada dasarnya, ada dua keuntungan yang diperoleh investor dengan membeli atau memiliki saham. Keuntungan pertama adalah saat pembagian laba yang diberikan oleh perusahaan, biasanya disebut dengan istilah *dividen*. Keuntungan yang kedua merupakan selisih antara harga beli dan harga jual. Keuntungan ini disebut dengan *capital gain*.

Di dalam aktivitas perdagangan saham sehari-hari, harga saham mengalami fluktuasi, baik berupa kenaikan maupun penurunan. Pembentukan harga saham terjadi karena adanya permintaan dan penawaran atas saham tersebut. Dengan kata lain, harga saham terbentuk oleh *supply* dan *demand* atas saham tersebut. Para investor biasanya melakukan analisis terhadap saham yang akan dibelinya, baik secara teknikal maupun secara fundamental [2].

B. Analisis Fundamental dan Analisis Teknikal

Analisis Teknikal adalah suatu teknik analisis yang digunakan untuk memprediksi harga saham dengan cara mempelajari suatu data transaksi jual beli yang lampau, terutama pergerakan harga dan volume. Dalam analisis teknikal, digunakan berbagai model dan dasar untuk memprediksi, misalnya dengan indeks kekuatan relatif, indeks pergerakan rata-rata ataupun dengan menganalisis pola grafik. Dalam penggunaannya, analisis teknikal lebih mengutamakan studi atas grafik harga.

Analisis fundamental adalah metode analisis yang didasarkan pada fundamental ekonomi suatu perusahaan. Teknik ini menitik beratkan pada rasio finansial dan kejadian-kejadian yang secara langsung maupun tidak langsung mempengaruhi kinerja keuangan perusahaan. Analisis fundamental lebih cocok untuk membuat keputusan dalam memilih saham perusahaan mana yang dibeli untuk jangka panjang. Analisis ini melibatkan banyak sekali variabel data yang harus dianalisis dan biasanya dilakukan dengan menggunakan rasio-rasio keuangan. Secara garis besar, rasio keuangan dibagi kedalam 5 kategori utama, antara lain rasio profitabilitas, rasio harga, rasio likuiditas, rasio daya ungkit dan rasio efisiensi [2].

C. Rasio Profitabilitas

Rasio profitabilitas merupakan rasio yang menggambarkan kemampuan perusahaan dalam mendapatkan laba melalui semua kemampuan dan

sumber yang ada seperti kegiatan penjualan, kas, modal, jumlah karyawan, jumlah cabang dan sebagainya. Dalam rasio profitabilitas, terdapat 6 rasio antara lain *gross profit margin*, *net profit margin*, *basic earning power*, *return on assets*, *return on equity* dan *earning per share*.

1. *Gross profit margin*

Merupakan rasio yang digunakan untuk mengukur efisiensi pengendalian harga pokok atau biaya produksinya. *Gross profit margin* dapat menilai kesehatan keuangan suatu perusahaan dengan mengungkapkan proporsi uang yang tersisa dari pendapatan setelah menghitung biaya pokok penjualan [3]. Dihitung dengan rumus :

$$\frac{(\text{Penjualan} - \text{harga pokok penjualan})}{\text{Penjualan}}$$

Metrik ini dapat digunakan untuk membandingkan perusahaan dengan pesaingnya. Biasanya sebuah perusahaan terlihat lebih tepat guna biasanya memiliki nilai *gross profit margin* yang tinggi

2. *Net profit margin*

Net profit margin merupakan rasio profitabilitas yang mengukur berapa banyak dari setiap uang penjualan perusahaan yang benar-benar membuat laba. Marjin ini diperoleh dengan mengukur laba bersih setelah pajak terhadap penjualan [3]. Dihitung dengan rumus:

$$\frac{\text{Laba bersih setelah pajak}}{\text{Penjualan}}$$

Net profit margin sangat berguna ketika membandingkan perusahaan di sektor industry yang sejenis. Perusahaan dengan *net profit margin* yang lebih tinggi menunjukkan bahwa perusahaan tersebut lebih memiliki kontrol yang baik lebih atas biaya produksinya.

3. *Basic eaning power*

Merupakan kemampuan bisnis sebuah perusahaan untuk menghasilkan keuntungan dari operasionalnya. *Basic earning power* adalah perbandingan laba sebelum pajak terhadap saham. Marjin ini digunakan untuk mendasari analisis apakah saham perusahaan tersebut layak untuk diinvestasikan [3]. Dihitung dengan rumus :

$$\frac{\text{Laba bersih sebelum pajak}}{\text{Total aktiva}}$$

4. *Return on assets*

Return on assets adalah indikator keuntungan perusahaan terhadap total asetnya. Marjin ini memperlihatkan seberapa efisien manajemen sebuah perusahaan dalam menggunakan asetnya

untuk menghasilkan laba. *Return on assets* merupakan perbandingan antara laba bersih setelah pajak dengan total aktiva [3]. Diperoleh dengan rumus :

$$\frac{\text{Laba bersih setelah pajak}}{\text{Total aktiva}}$$

5. *Return on equity*

Merupakan rasio pengukuran dari penghasilan yang tersedia bagi para pemilik modal atas perusahaan yang mereka investasikan dalam perusahaan. *Return on equity* memperlihatkan laba bersih yang diperoleh para pemegang sahamnya dalam bentuk persentase [3]. Diperoleh dengan rumus :

$$\frac{\text{Laba bersih setelah pajak}}{\text{Ekuitas}}$$

6. *Earning per share*

Earning per share menunjukkan seberapa besar kemampuan perlembar saham dalam menghasilkan laba [3]. Diperoleh dengan rumus :

$$\frac{\text{Laba bersih setelah pajak} - \text{dividen saham perferen}}{\text{Jumlah saham yang beredar}}$$

D. *Extraction, Transformation, Loading (ETL)*

Pendataan pada *Data warehouse* harus dibuat teratur agar dapat memfasilitasi analisis bisnis. Untuk melakukan hal ini, data dari satu atau lebih sistem operasional perlu di ekstrasi dan disalin kedalam *data warehouse*, proses ini disebut *ETL* (*extraction, transformation, loading*). Secara keseluruhan, *ETL* sering didefinisikan sebagai rekonsiliasi. *ETL* dilakukan pada populasi pertama gudang data dan akan dilakukan kembali secara berkala pada proses *update* gudang data. *ETL* terdiri dari empat fase terpisah : *extraction, cleaning, transformation* dan *loading* [1].

1. *Extraction*

Selama *extraction*, data yang diinginkan diidentifikasi dan diambil dari berbagai sumber, termasuk sistem database dan aplikasi. Data yang relevan diperoleh dari sumber-sumber pada fase *extraction*. *Static extraction* dilakukan ketika gudang data dipopulasikan pertama kali. *Incremental extraction* digunakan untuk *update* gudang data secara berkala.

2. *Cleaning*

Fase *cleaning* sangat penting dalam sistem *data warehouse* karena akan meningkatkan kualitas data, yang pada umumnya cukup buruk pada sumber asalnya. Beberapa faktor yang meliputi kesalahan dan inkonsistensi yang paling sering terjadi membuat data “kotor” :

- Penggunaan field yang tidak sesuai
- Duplikasi data
- Masukkan yang salah
- Nilai-nilai tidak konsisten pada satu entitas karena perbedaan aturan
- Nilai-nilai yang tidak konsisten untuk satu entitas individu karena salah ketik

Fitur inti data *cleaning* yang ditemukan dalam perkakas *ETL* adalah pembedahan (*rectification*) dan homogenisasi (*homogenization*). Mereka menggunakan kamus-kamus spesifik untuk melakukan pembedahan salah ketik dan untuk mengenali sinonim dan juga sebagai *rule-based cleansing* untuk memaksa aturan spesifik domain dan mendefinisikan asosiasi antar nilai yang sesuai.

3. *Transformation*

Transformasi adalah inti dari fase rekonsiliasi. Pada tahap ini dilakukan konversi data. Proses transformasi ditujukan pada konversi dan normalisasi pada format penyimpanan untuk membuat data seragam. Transformasi ini juga bertujuan menyatukan field-field yang memiliki nilai yang sama pada berbagai sumber. Ketika mempopulasikan *data warehouse*, normalisasi digantikan dengan denormalisasi karena *data warehouse* secara khusus di denormalisasi dan dibutuhkan agregasi untuk meringkas data dengan cepat.

4. *Loading*

Pemuatan ke dalam *data warehouse* adalah tahap terakhir untuk dilakukan. Pemuatan dilakukan dengan cara *refresh*, yaitu gudang data ditulis ulang secara keseluruhan dan *update*, yaitu hanya perubahan yang dilakukan pada sumber data yang akan ditambahkan ke dalam gudang data.

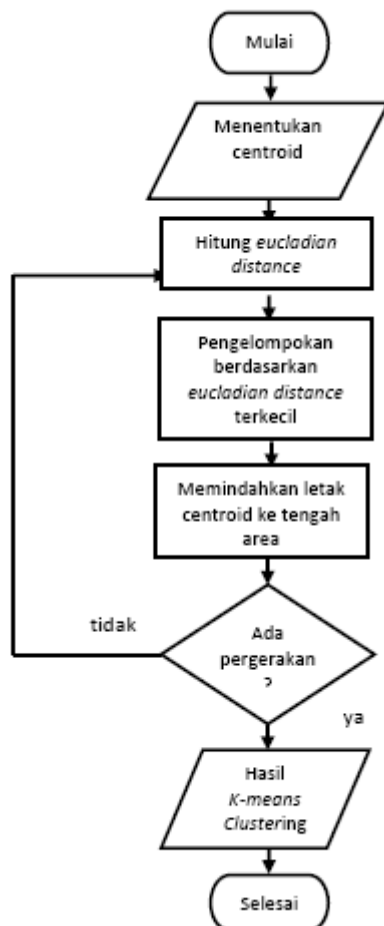
E. *K-Means Clustering*

K-means adalah salah satu bentuk pengelompokan yang paling sederhana. Prosedurnya sederhana dan mudah untuk mengklasifikasikan data yang diberikan melalui sejumlah *cluster*. Penentuan *centroid* dilakukan dengan cara mengambil data pertama sebagai *centroid* pertama, data kedua sebagai *centroid* kedua, dan seterusnya hingga jumlah *centroid* yang diperlukan. Langkah berikutnya adalah dengan menghitung jarak dari titik yang akan di *cluster* ke setiap *centroid* yang ada dan dikelompokkan sesuai dengan jarak terdekat kepada *centroid*-nya. Bila semua titik sudah masuk kedalam pengelompokan maka langkah pertama selesai. Kemudian langkah berikutnya, kita perlu menghitung kembali *k-centroid* baru sebagai *barycenters* dari kelompok yang dihasilkan. Setelah memiliki *k-centroid* yang baru, pengelompokan di uji kembali terhadap *k-centroid*. Penghitungan *k-centroid*

yang baru dan pengelompokannya dilakukan berulang hingga k -centroid tidak bergerak lagi. Algoritmanya adalah :

$$J = \sum_{j=1}^k \sum_{i=1}^n \|x_i^{(j)} - c_j\|^2$$

Dimana $\|x_i^{(j)} - c_j\|^2$ adalah ukuran jarak yang dipilih antara titik data $x_i^{(j)}$ dan pusat cluster c_j [4].



Gambar 1. Flowchart Algoritma *K-Means Clustering*.

Dari gambar 1 di atas dapat dilihat bagaimana flowchart untuk *K-Means* bekerja. Akan terjadi proses perulangan sampai dengan pergerakan yang dihasilkan berhenti. Implementasi *K-Means clustering* ini telah digunakan untuk beberapa penelitian, diantaranya untuk melakukan pengelompokan terhadap konsumen pada sisi tipe dari konsumen, loyalitas konsumen [4]. Sebagai dasar terdapat 5 kelompok yang akan digunakan dalam rancangan awal. Dan sebagai hasil akhirnya, algoritman *K-Means* digabungkan dengan algoritma C.4.5 untuk menerapkan loyalitas konsumen.

Algoritma *K-Means* juga diterapkan oleh Alghamdi dan Nor (2014) untuk meningkatkan representasi dokumen dengan menerapkan pula algoritma text mining. Pendekatan ini juga dengan menerapkan pendekatan *k-Means clustering* yang dipadukan dengan Vektor Bayesian. Sebagai hasil akhir dari perancangan tersebut, menunjukkan bahwa algoritma ini dapat digunakan untuk meningkatkan performa dari penerapan algoritma *k-Means*. Dan hasil akhir dari penelitian ini juga menerapkan bahwa didapatkan 95% lebih dekat hasilnya di dibandingkan dengan algoritma *k-Means* yang konvensional.

Implementasi algoritma *K-Means* juga digunakan untuk menerapkan rancangan terhadap harga saham dan harga komoditas [5]. Santoso, Albertus, dan Eduard (2013) menerapkan algoritma *K-Means* yang dikombinasikan dan digabung dengan Principal Component Analysis dan Artificial Neural Networks. Algoritma *K-Means* digunakan untuk mengetahui relasi antara harga saham dengan harga komoditas. Adapun harga komoditas yang digunakan adalah harga untuk perunggu, perak, minyak, gas alam, dan emas. Setelah itu, algoritma ini dikembangkan untuk mendapatkan keunikan dan identifikasi terhadap masing – masing cluster dengan menggunakan metode Principal Component Analysis. Dari hasil Principal Component Analysis tersebut di dapatkan identifikasi dari masing – masing cluster. [5]. Untuk menguji coba apakah data training (pelatihan), maka metode Backpropagation dari Jaringan Syaraf Tiruan diterapkan. Sebagai hasilnya nampak bahwa menerapkan metode *K-Means* dengan Principal Component Analysis merupakan salah satu langkah yang cukup efisien [5].

K-Means clustering juga dapat digunakan untuk melakukan pengelompokan terhadap kerentanan manusia terhadap paparan bahan – bahan kimia yang berbahaya di dalam kota industri di China [6]. Shi dan Weihua (2013) menerapkan algoritma *K-Means* yang berbasis pada algoritma genetik. Dari hasil pengelompokan tersebut, didapatkan suatu cluster yang optimal yang terbagi menjadi 6 pola. Cluster 1 dan 4 membutuhkan perhatian yang lebih tinggi dibandingkan cluster yang lainnya [6].

III. RANCANGAN SISTEM

A. Analisa Data

Pengumpulan data dilakukan dengan membaca laporan keuangan tahunan yang telah di audit. Laporan diperoleh dalam bentuk *softcopy*, diunduh dari www.idx.co.id. Dari laporan tersebut, didapatkan berbagai data yang diperlukan dalam menghitung rasio profitabilitas. Data yang digunakan untuk menghitung rasio profitabilitas antara lain nilai penjualan, harga pokok penjualan, laba bersih setelah pajak, laba bersih sebelum pajak, total aktiva, deviden saham, ekuitas dan total saham beredar. Data ini akan dipergunakan untuk menghitung rasio profitabilitas.

Langkah awal yang dilakukan pada saat pengumpulan data adalah mencari data keuangan dari 68 emiten dalam 1 periode. Laporan tahunan setiap perusahaan dalam periode yang sama dibaca dan dari laporan tersebut diperoleh data- untuk menghitung rasio profitabilitas. Untuk mempermudah penyimpanan data dan penghitungan rasio profitabilitas, data tersebut di olah ke dalam aplikasi *Microsoft Excel*.

Sebagai contoh, dilakukan terhadap laporan keuangan PT.Akasha Wira Internasional.Tbk (ADES) periode 2012. Dari gambar 2 dapat kita peroleh nilai-nilai penjualan dari PT ADES sebagai berikut :

- Penjualan = Rp 476.638.000.000
- Harga pokok penjualan = Rp 204.736.000.000
- Laba bersih setelah pajak = Rp 83.376.000.000
- Laba bersih sebelum pajak = Rp 76.631.000.000
- Total aktiva = Rp 389.094.000.000
- Deviden saham perferen = Rp 141
- Ekuitas = Rp 209.122.000.000
- Total saham beredar = 589.896.800 lembar

PT AKASHA WIRA INTERNATIONAL Tbk

LAPORAN LABA RUGI KOMPREHENSIF UNTUK TAHUN YANG BERAKHIR PADA TANGGAL-TANGGAL 31 DESEMBER 2012 DAN 2011 (Dinyatakan dalam Jutaan Rupiah, kecuali Dinyatakan Lain)

	Catatan/ Notes	2012		2011	
E K U I T A S					
Modal Saham					
Modal Dasar - 2.359.587.200 saham					
Modal Ditempatkan dan Disetor Penuh - 589.896.800 saham dengan nilai nominal Rp 1.000 (dalam angka penuh) per saham		589.897		589.897	
Tambahan Modal Disetor	17	74.920		49.052	
Saldo Laba (Rugi) :	18	(460.763)		(518.271)	
- Dicadangkan	19	209.122		125.746	
- Belum Dicadangkan					
Jumlah Ekuitas		389.094		316.048	
JUMLAH UABILITAS DAN EKUITAS					
PENJUALAN BERSIH	2n,20&30	476.638		299.409	
BEBAN POKOK PENJUALAN	2n,21,30&33	(204.736)		(184.925)	
LABA KOTOR		271.902		114.484	
Beban Penjualan	2n,22&30	(116.795)		(61.999)	
Beban Umum dan Administrasi	2n,23&30	(68.699)		(26.186)	
Beban Lain-lain	24 & 30	(1.752)		(741)	
Penghasilan Lain-lain	24 & 30	6.790		25.852	
Penghasilan Keuangan	25,30&33	302		191	
Beban Keuangan	26,30&33	(15.117)		(21.994)	
LABA SEBELUM PAJAK PENGHASILAN		76.631		29.627	
MANFAAT (BEBAN) PAJAK PENGHASILAN	2o,2r&27	6.745		(3.759)	
LABA TAHUN BERJALAN		83.376		25.868	
Pendapatan Komprehensif Lain		-		-	
JUMLAH PENDAPATAN KOMPREHENSIF TAHUN BERJALAN		83.376		25.868	
LABA BERSIH PER SAHAM (dalam Angka Penuh)	2p	141		44	
RATA-RATA TERTIMBANG JUMLAH SAHAM BEREDAR/DITEMPATKAN (dalam Angka Penuh)		589.896.800		589.896.800	

*) Direklasifikasi, lihat Catatan 33/ Reclassified, refer to Note 33

Gambar 2. Laporan Keuangan PT.Akasha Wira Internasional.Tbk(ADES) periode 2012.

Kemudian dilakukan terhadap 67 laporan keuangan perusahaan lainnya pada periode yang sama yaitu tahun 2012. Data yang diperoleh, disimpan kedalam file *Microsoft excel (.xls)* yang nantinya akan diproses untuk menghitung rasio profitabilitas. Data periode 2012 dapat dilihat pada gambar 3.

1	Column	Column2	Column3	Column4	Column5
2	ADES	IDR 476,638,000,000.00	IDR 204,736,000,000.00	IDR 83,376,000,000.00	IDR 76,631,000,000.00
3	ADRO	\$ 3,722,489.00	\$ 2,679,967.00	\$ 383,307.00	\$ 713,724.00
4	AISA	IDR 2,747,623,000,000.00	IDR 2,142,377,000,000.00	IDR 253,664,000,000.00	IDR 324,465,000,000.00
5	ALTO	IDR 233,675,793,803.00	IDR 186,709,508,830.00	IDR 16,167,317,065.00	IDR 21,955,756,592.00
6	ANTM	IDR 10,449,885,512,000.00	IDR 8,427,157,554,000.00	IDR 2,989,024,589,000.00	IDR 3,895,495,061,000.00
7	ARII	\$ 97,240,000.00	\$ 80,674,000.00	\$ (11,150,000.00)	\$ (14,040,000.00)
8	ARTI	IDR 449,486,392,992.00	IDR 278,254,523,220.00	IDR 51,857,031,148.00	IDR 56,735,886,400.00
9	ATPK	IDR 181,494,610,000.00	IDR 168,874,180,000.00	IDR (16,740,643,000.00)	IDR (16,740,643,000.00)
10	BIPI	IDR 357,108,000,000.00	IDR 231,891,000,000.00	IDR (144,259,000,000.00)	IDR 8,916,000,000.00
11	BORN	\$ 645,950,403.00	\$ (416,127,790.00)	\$ (550,456,125.00)	\$ (520,247,602.00)
12	BRAU	\$ 1,531,063.00	\$ 1,024,380.00	\$ (179,510.00)	\$ 1,485,000.00
13	BSSR	\$ 108,968,528.00	\$ 65,054,321.00	\$ 9,783,589.00	\$ 12,911,035.00
14	BUMI	\$ 2,768,468,676.00	\$ 2,018,139,069.00	\$ (655,424,101.00)	\$ (596,644,341.00)
15	BYAN	\$ 1,422,880,281.00	\$ 1,164,147,320.00	\$ 62,449,087.00	\$ 78,548,998.00
16	CEKA	IDR 1,123,519,657,631.00	IDR 955,696,220,119.00	IDR 58,344,237,476.00	IDR 83,714,325,804.00
17	CITA	IDR 2,608,869,416,336.00	IDR 1,456,967,440,962.00	IDR 245,081,455,043.00	IDR 331,252,165,417.00
18	CKRA	IDR 24,523,649,769.00	IDR 19,929,460,342.00	IDR (3,957,879,273.00)	IDR (4,316,901,574.00)
19	CTTH	IDR 161,783,288,701.00	IDR 102,976,934,013.00	IDR 2,759,299,965.00	IDR 3,198,037,734.00
20	DAVO	IDR 1,210,836,618,300.00	IDR 1,248,327,670,974.00	IDR (2,695,748,296,504.00)	IDR 2,876,207,469,151.00
21	DEWA	\$ 222,028,647.00	\$ 232,668,420.00	\$ (50,647,320.00)	\$ (53,447,724.00)
22	DKFT	IDR 848,500,274,577.00	IDR 481,733,214,139.00	IDR 303,447,766,556.00	IDR 393,668,574,628.00
23	DLTA	IDR 1,719,814,548,000.00	IDR 719,951,793,000.00	IDR 213,421,077,000.00	IDR 287,505,070,000.00
24	DOID	\$ 843,254,769.00	\$ 740,246,140.00	\$ (17,368,064.00)	\$ (17,698,665.00)
25	DVLA	IDR 1,087,379,869,000.00	IDR 436,269,979,000.00	IDR 148,909,089,000.00	IDR 204,477,046,000.00

1	Column	Column6	Column7	Column8	Column9
2	ADES	IDR 389,094,000,000.00	IDR 209,122,000,000.00	IDR 141.00	589,896,800
3	ADRO	\$ 6,692,256.00	\$ 2,995,054.00	\$ 0.01	31,985,962.000
4	AISA	IDR 3,867,576,000,000.00	IDR 2,033,453,000,000.00	IDR 72.18	2,926,000,000.00
5	ALTO	IDR 324,619,954,340.00	IDR 188,921,883,796.00	IDR 10.43	1,550,000.000
6	ANTM	IDR 19,708,540,946,000.00	IDR 12,832,316,056,000.00	IDR 314.06	9,538,459,750
7	ARII	\$ 299,105,000.00	\$ 144,306,000.00	\$ (0.00)	91,999
8	ARTI	IDR 1,432,238,854,122.00	IDR 855,759,068,630.00	IDR 33.07	1,568,000.000
9	ATPK	IDR 150,829,602,000.00	IDR 43,824,058,000.00	IDR (18.31)	914,324,669
10	BIPI	IDR 4,479,157,000,000.00	IDR 3,723,716,000,000.00	IDR 0.28	35,218,521,254
11	BORN	\$ 2,061,533,012.00	\$ 352,581,276.00	\$ (0.03)	17,693,000.000
12	BRAU	\$ 2,148,128.00	\$ 242,272.00	\$ (0.01)	34,900,000.000
13	BSSR	\$ 139,309,783.00	\$ 82,468,253.00	\$ 0.00	2,616,500.000
14	BUMI	\$ 7,157,394,649.00	\$ 468,159,613.00	\$ (31.16)	473,212,607
15	BYAN	\$ 1,909,104,988.00	\$ 707,691,110.00	\$ 0.02	3,333,333,500
16	CEKA	IDR 1,027,692,718,504.00	IDR 463,402,986,308.00	IDR 196.00	297,500.000
17	CITA	IDR 1,968,579,105,393.00	IDR 1,134,652,334,465.00	IDR 58.00	3,370,734,900
18	CKRA	IDR 1,213,229,329,031.00	IDR 1,172,825,729,969.00	IDR (0.69)	13,448,400.000
19	CTTH	IDR 261,438,526,210.00	IDR 78,751,784,810.00	IDR 2.24	1,230,839,821
20	DAVO	IDR 2,510,202,205,485.00	IDR 2,126,405,553,626.00	IDR 217.00	12,403,711,320
21	DEWA	\$ 365,758,029.00	\$ 222,107,938.00	\$ (2.37)	21,853,733,792
22	DKFT	IDR 1,535,650,131,037.00	IDR 1,386,445,565,793.00	IDR 55.00	5,612,355,730
23	DLTA	IDR 745,306,835,000.00	IDR 598,211,513,000.00	IDR 12.99	16,013,181
24	DOID	\$ 1,159,770,820.00	\$ 89,506,873.00	\$ 0.00	8,148,494,232
25	DVLA	IDR 1,074,691,476,000.00	IDR 436,269,979,000.00	IDR 133.00	1,120,000,000

Gambar 3. Gambar Data Laporan Keuangan 2012 pada Microsoft Excel.

Pengolahan data pada file ini dilakukan dengan menggabungkan keseluruhan laporan keuangan setiap emiten pada periode tahun yang sama. Setelah data diperoleh, maka dilakukan perhitungan pada data tersebut untuk mendapatkan rasio profitabilitas.

Sebelum melakukan *clustering* menggunakan *K-means*, dilakukan pembersihan data. Fase ini sangat penting dalam sistem gudang data karena meningkatkan kualitas data. Tujuan dari *fase cleaning* yaitu pembersihan dan homogenisasi data. Data yang akan diproses menggunakan kurs rupiah (IDR), maka pada data yang menggunakan kurs dolar amerika (\$) akan dilakukan konversi ke dalam kurs rupiah. Nilai tukar yang

digunakan adalah nilai tukar yang tertera dalam laporan keuangan setiap emiten.

Setelah dilakukan konversi kurs, langkah berikutnya adalah melakukan pengolahan data. Data yang dibutuhkan adalah data *profit* (keuntungan), sehingga semua data yang bernilai negatif (kerugian) tidak diikuti sertakan dalam *clustering*. Setelah menyingkirkan data yang bernilai negatif, dilakukan penghitungan untuk mendapatkan nilai-nilai rasio profitabilitas, sebagai berikut :

1. $Gross\ profit\ margin = \frac{(476.638.000.000 - 204.736.000.000)}{476.638.000.000} = 0,57 = 57\%$
2. $Net\ profit\ margin = \frac{83.376.000.000}{476.638.000.000} = 0,17 = 17\%$
3. $Basic\ earning\ power = \frac{389.094.000.000}{83.376.000.000} = 0,2 = 20\%$
4. $Return\ on\ assets = \frac{389.094.000.000}{83.376.000.000} = 0,21 = 21\%$
5. $Return\ on\ equity = \frac{209.122.000.000}{(83.376.000.000 - 141)} = 0,4 = 40\%$
6. $Earning\ per\ share = \frac{(83.376.000.000 - 141)}{589.896.800} = 141,34$

Kemudian, perhitungan dilakukan terhadap 67 perusahaan lainnya pada periode 2012. Data yang diperoleh kemudian dimasukkan ke dalam microsoft excel lagi. Setelah memperoleh rasio profitabilitas emiten-emiten lainnya, cleaning dilakukan sekali lagi. Proses cleaning disini bertujuan untuk menghilangkan data yang nilainya lebih dari 100%. Setelah data yang dibutuhkan sesuai, fase terakhir adalah melakukan *k-means clustering*.

IV. IMPLEMENTASI SISTEM

Pada proses ini, data yang diinputkan merupakan data yang akan diproses melalui metode *K-means Clustering*. Proses input data dilakukan melalui form setup data seperti pada gambar 4.

Untuk memasukkan data yang akan di proses, pengguna harus mengakses tombol *browse*, kemudian memilih file Microsoft office (.xls) yang akan di proses. Setelah file dipilih, data akan masuk kedalam form setup data, kemudian pengguna dapat memproses untuk menghitung rasio profitabilitas dengan mengakses tombol Rasio. Pada proses ini, sistem akan menghitung 6 rasio profitabilitas, yaitu *gross profit margin*, *net profit margin*, *basic earning power*, *return on assets*, *return on equity* dan *earning per share*.

Setelah mendapatkan hasil perhitungan, pengguna dapat melanjutkan ke proses berikutnya untuk melihat penyajian dalam diagram scatter. Sistem secara otomatis akan menampilkan data dengan sumbu x bernilai *gross profit margin* dan sumbu y bernilai *net profit margin*. Selanjutnya proses *clustering* dilakukan dengan mengkases tombol *cluster*. Setelah dilakukan ujicoba pada periode 2009 dan mencoba menggunakan 2 centroid hingga 5 centroid.

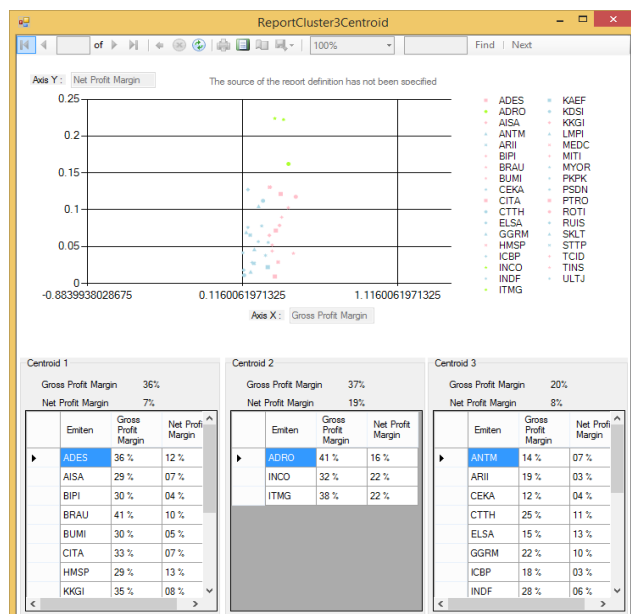
Emiten	Penjualan	Harga Pokok Penjualan	Labas Bersih Setelah Pajak	Labas Bersih Sebelum Pajak
ADES	476638000000	204736000000	83376000000	76631000000
ADRO	3722489	2679967	383307	713724
AISA	2747623000000	2142377000000	253664000000	324465000000
ALTO	233675793803	186709508830	16167317065	21955756592
ANTM	10449885512000	8427157554000	2989024589000	3895495061000
ARII	97240000	80674000	-11150000	-14040000
ARTI	449486392992	278254523220	51857031148	56735886400
ATPK	181494610000	168874180000	-16740643000	-16740643000
BIPI	357108000000	231891000000	-144259000000	8916000000
BORN	645950403	-416127790	-550456125	-520247602
BRAU	1531063	1024380	-179510	1485000
BSSR	108968528	65054321	9783589	12911035
BUMI	2768488676	2018139069	-655424101	-596644341
BYAN	1422880281	1164147320	62449087	78548998
CEKA	1123519657631	955696220119	58344237476	83714325804
CITA	2608869416336	1456967440962	245081455043	331252165417

Emiten	Total Aktiva	Ekuitas	Deviden Saham Preferen	Total Saham Beredar
ADES	389094000000	209122000000	141	589896800
ADRO	6692256	2995054	0.012	31985962000
AISA	3867576000000	2033453000000	72.18	2926000000
ALTO	324619954340	188921883796	10.43	1550000000
ANTM	19708540946000	12832316056000	314.06	9538459750
ARII	299105000	144306000	-0.0036	91999
ARTI	1432238854122	855759068630	33.07	1568000000
ATPK	150829602000	43824058000	-18.31	914324669
BIPI	4479157000000	3723716000000	0.28	35218521254
BORN	2061533012	352581276	-0.032	17693000000
BRAU	2148128	242272	-0.0052	34900000000
BSSR	139309783	82468253	0.004	2616500000
BUMI	7157394649	468159613	-31.16	473212607
BYAN	1909104988	707691110	0.02	333333500
CEKA	1027692718504	463402986308	196	2975000000
CITA	1968579105393	1134652334465	58	3370734900

Gambar 4. Form Setup Data

Gambar 5 memperlihatkan hasil clustering menggunakan 3 centroid., hasilnya di masukkan kedalam tabel untuk melihat perbandingannya. Tabel 1 memperlihatkan hasil akhir ujicoba periode 2009.

Dari percobaan maka dapat disimpulkan bahwa semakin banyak jumlah centroidnya, maka hasil pengelompokan terhadap centroidnya semakin jelas. Saat jumlah centroid 2, kelompok centroid 1 dan centroid 2 terdapat banyak kemiripan, range antara batas bawah dan batas bawahnya tercampur. Ketika jumlah centroid diubah menjadi 3 hingga menjadi 5, perbedaan di setiap kelompoknya makin terlihat. Batas bawah antara kelompok yang satu dan kelompok yang lainnya makin spesifik.



Gambar 5. Clustering menggunakan 3 centroid

Tabel 1. Tabel Hasil Clustering Periode 2009

		K = 2 (%)	K = 3 (%)	K = 4 (%)	K = 5 (%)
Centroid 1	Gross profit margin	12 – 32	29 – 46	29 – 44	30 – 44
	Net profit margin	1 – 13	1 – 13	1 – 13	1 – 12
Centroid 2	Gross profit margin	32 – 44	32 – 41	38 – 41	41 – 46
	Net profit margin	3 – 22	16 – 22	16 – 22	12 – 16
Centroid 3	Gross profit margin	-	12 – 32	32	32 – 38
	Net profit margin	-	1 – 13	22	22
Centroid 4	Gross profit margin	-	-	12 – 25	12 – 22
	Net profit margin	-	-	1 – 13	1 – 13
Centroid 5	Gross profit margin	-	-	-	22 – 29
	Net profit margin	-	-	-	2 – 13

V. PENUTUP

A. Kesimpulan

Berdasarkan hasil coba terhadap laporan keuangan selama lima tahun, didapat kesimpulan:

1. Berhasil membangun dan menghasilkan sistem berbasis desktop yang dapat mengelompokkan rasio profitabilitas pada tahun yang sama dengan kemiripan pada 2 parameter yang dipilih.
2. Metode *K-Means* dapat menghasilkan cluster rasio profitabilitas dan melihat kemiripan nilainya dalam setiap clusternya.
3. Semakin besar nilai K pada *K-Means*, maka semakin menunjukkan kelompok rasio profitabilitas dan nilai kelompok makin mendekati kemiripan.

B. Saran

Untuk penelitian lebih lanjut, dapat menambahkan metode *K-medoids* ke dalam aplikasi. Centroid akan langsung diwakili oleh satu emiten tertentu sehingga cluster terhadap rasio profitabilitas dapat diperoleh lebih khusus dan lebih mengelompok. Sedangkan pada tabel hasil akhirnya, emiten yang analisis dapat dibedakan dari sektor mana, sehingga pada hasil pengelompokannya dapat dianalisis apakah sektor mempengaruhi pengelompokan.

DAFTAR PUSTAKA

- [1] Alghamdi, H. M., Nor Shahriza. (2014). Improved Text Clustering using K-Mean Bayesian Vectoriser. *Journal of Information and Knowledge Management*, Vol 13, No 3, 1-10.
- [2] Ali El-Sappagh, S. H., Ahmed Hendawi, M. A., & El Bastawissy, H. A. (2011). A Proposed Model for Data Warehouse ETL Processes. *Journal of King Saud University - Computer and Information Sciences*, 91-104
- [3] Bajkowski, J. (2010). Financial Ratio Analysis: Putting the Numbers to Work. *The American Association of Individual Investors*, 3-7.
- [4] Bunnak, P., Sotarat Thammaboosadee, Supaporn Kiattisin. (2015). Applying Data Mining Techniques and Extended RFM Model in Customer Loyalty Measurement. *Journal of Advances in Information Technology*. Vol. 6, No. 4, November, 238 – 242.
- [5] Santoso, H. B., Albertus Joko Santoso, Eduard Rusdianto. (2013). Hybrid Clustering Method for Stock Price and Commodity Price. *International Journal of Science and Advanced Technology*. Vol. 3, No. 7, July, 21- 29.
- [6] Shi, W., Weihua Zeng. (2013). Genetic K-Means Clustering Approach for Mapping Human Vulnerability to Chemical Hazards in the Industrialized City: A Case Study of Shanghai, China. *International Journal of Environmental Research and Public Health*. Vol. 10, 2578 – 2595.