

THESIS

**KLASTERISASI HARGA SAHAM DAN KOMODITAS
MENGUNAKAN METODE HYBRID
KLASTERISASI**



Halim Budi Santoso

125301830

PROGRAM STUDI MAGISTER TEKNIK INFORMATIKA
PROGRAM PASCASARJANA
UNIVERSITAS ATMA JAYA YOGYAKARTA
2013



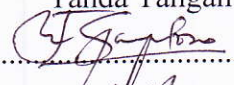
UNIVERSITAS ATMA JAYA YOGYAKARTA

PROGRAM PASCA SARJANA

PROGRAM STUDI MAGISTER TEKNIK INFORMATIKA

PERSETUJUAN THESIS

Nama : Halim Budi Santoso
Nomor Mahasiswa : 125301830 / PS / MTF
Konsentrasi : Enterprise Information System
Judul Thesis : Klasterisasi Harga Saham dan Komoditas
Menggunakan Metode Hybrid Klasterisasi

Nama Pembimbing	Tanggal	Tanda Tangan
Dr. Ir. Albertus Joko Santoso, MT	29 October 2013	
Eduard Rusdianto, ST, MT	25/10/2013	



UNIVERSITAS ATMA JAYA YOGYAKARTA

PROGRAM PASCA SARJANA

PROGRAM STUDI MAGISTER TEKNIK INFORMATIKA

PENGESAHAN THESIS

Nama : Halim Budi Santoso
 Nomor Mahasiswa : 125301830 / PS / MTF
 Konsentrasi : Enterprise Information System
 Judul Thesis : Klasterisasi Harga Saham dan Komoditas
 Menggunakan Metode Hybrid Klasterisasi

Nama Penguji	Tanggal	Tanda Tangan
Dr. Ir. Albertus Joko Santoso, MT (Ketua)	29 Oktober 2013	
Eduard Rusdianto, ST, MT (Anggota)	25/10/2013	
F. Sapty Rahayu, S.T., M.Kom. (Anggota)	29/10/2013	

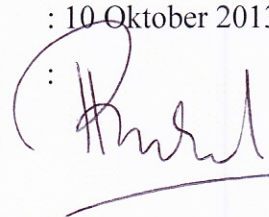
Ketua Program Studi

Dra. Ernawati, MT

Halaman Pernyataan

Thesis ini disusun sebagai hasil karya sendiri. Semua sumber yang digunakan pada thesis ini telah dikutip ataupun dirujuk dengan baik dan telah saya nyatakan dengan benar

Nama : Halim Budi Santoso
NIM : 125301830 / PS / MTF
Program Studi : Magister Teknik Informatika
Tanggal : 10 Oktober 2013
Tanda Tangan :



Intisari

Volatilitas harga saham adalah salah satu resiko dalam berinvestasi. Investor harus memiliki perkembangan informasi tentang harga saham. Dari perkembangan harga tersebut, investor dapat memutuskan apakah akan membeli, menjual, atau menahannya. Harga saham ini dipengaruhi oleh banyak faktor. Salah satu indikasi yang sering digunakan adalah harga komoditas. Apabila harga komoditas mengalami perubahan yang signifikan, apakah perubahan tersebut diikuti dengan perubahan harga saham. Oleh karena itu, perlu adanya suatu studi kasus untuk mengetahui bagaimana suatu pergerakan harga komoditas terhadap pergerakan harga saham. Penelitian ini dilakukan untuk melakukan pembelajaran mengenai penerapan metode K-Means, *Principal Component Analysis*, dan Jaringan Syaraf Tiruan terhadap harga saham dan harga komoditas.

Metode K-Means diterapkan untuk mengelompokkan harga saham dan harga komoditas. Diharapkan dengan mengelompokkan harga saham tersebut ke dalam suatu kluster dapat mengetahui dampak harga komoditas terhadap suatu sektor tertentu. Setelah kluster terbentuk, metode *Principal Component Analysis* diterapkan untuk mengurangi variabel – variabel yang tidak berkorelasi dalam suatu kluster. Selain itu, metode Jaringan Syaraf Tiruan dengan menggunakan Backpropagation diterapkan untuk membentuk jaringan syaraf dari tiap kluster.

Penerapan metode K-Means pada pergerakan harga saham dan harga komoditas mendapatkan kluster yang berisi informasi pergerakan harga komoditas dan korelasinya terhadap harga saham sektor tertentu. Penerapan metode *Principal Component Analysis* dapat digunakan untuk mereduksi variabel – variabel yang tidak berkorelasi. Dengan menerapkan metode *Principal Component Analysis* juga mengakibatkan kinerja Jaringan Syaraf Tiruan yang terbentuk menjadi bekerja semakin baik.

Kata Kunci : Klasterisasi K-Means, *Principal Component Analysis*, Jaringan Syaraf Tiruan, Harga saham, Harga Komoditas

Abstract

Stock price volatility is one of the investment risks. Investor should have updated information. Based on those information, Investor will decide whether want to buy, sell, or hold the stock. Stock price volatility is influenced by some factors related with the global economy. One of the them is commodity price. If the commodity price were to drop sharply, is there any impact to the specific company. Based on this question, this study is conducted. This study is conducted to test the K-Means Clustering, Principal Component Analysis, and Backpropagation Neural Network for stock price and commodity price.

K-Means Clustering is conducted to group the similar movement of the stock price and commodity price. After the cluster is established, Principal Component Analysis is implemented for every cluster. Besides, Artificial Neural Network is implemented to every cluster using Backpropagation Neural Network. There will be comparison study about the performace of the Artificial Neural Network for compact cluster and uncompact cluster.

Implementation of the K-Means Clustering will make ten cluster of commodity price and the effect to the spesific sector. This study also found that Principal Component Analysis can reduce the dimension of every cluster. Principal Component Analysis also can improve the performace by reducing the Mean Squared Error of every network for every cluster.

Keywords : K-Means Clustering, Principal Component Analysis, Backpropagation Neural Network, Stock Price, Commodity Price.

Kata Pengantar

Puji syukur kepada Tuhan Yesus yang sudah memberikan kekuatan dan berkatNya selalu sehingga penelitian ini dapat terselesaikan di tengah kesibukan dan rutinitas sehari – hari.

Penulis juga mengucapkan syukur kepada orang tua yang selalu memberikan doa yang terbaik sehingga penelitian ini dapat diselesaikan dengan baik. Penulis tak lupa mengucapkan terima kasih atas bantuan dari Yayasan Arsari Djojohadikusumo atas beasiswa yang diberikan sehingga penulis dapat menyelesaikan studi dan penelitian ini dengan baik.

Bagi pak Albertus Joko Santoso dan pak Eduard Rusdianto, penulis juga mengucapkan terima kasih atas bantuan dan bimbingan selama ini. Penulis merasa sangat berterima kasih atas nasehat yang diberikan untuk menyempurnakan penelitian ini.

Semoga dengan adanya penelitian ini dapat memberikan kontribusi terhadap penerapan data mining untuk faktor keuangan, khususnya di dalam bidang investasi. Penulis juga menyadari bahwa penelitian ini masih jauh dari sempurna. Oleh karena itu, penulis terbuka terhadap saran dan masukan bagi kontribusi penelitian ini.

Yogyakarta, 4 Oktober 2013

Halim Budi Santoso

Daftar Isi

Halaman Judul	i
Halaman Persetujuan	ii
Halaman Pengesahan	iii
Halaman Pernyataan Keaslian	iv
Intisari	v
Abstract	vi
Kata Pengantar	vii
Daftar Isi	viii
Daftar Tabel	x
Daftar Gambar	xi
BAB I Pendahuluan	1
A. Latar Belakang Masalah	1
B. Rumusan Masalah	3
C. Batasan Masalah	4
D. Tujuan Penelitian	4
E. Manfaat Penelitian	5
1 Kontribusi Praktis	5
2 Kontribusi Teoritis	5
F. Sistematika Penulisan	5
BAB II Landasan Teori	7
A. Tinjauan Pustaka	7
B. Landasan Teori	14
1 <i>Simple Moving Average (SMA)</i>	14
2 Penambangan data	15
3 Analisa Klaster	16
4 <i>Similarity dan Disimilarity</i>	18
5 K-Means Klasterisasi	19
6 <i>Principal Component Analysis</i>	20
7 Jaringan Syaraf Tiruan	22
BAB III Metodologi Penelitian	25
A. Bahan dan Materi	25
1 Dataset dan Statistik Deskriptif	27
B. Alat Pendukung Penelitian	30
C. Langkah Penelitian	30
1 Pra Penelitian	30

2	Pembersihan Data (<i>Data Cleaning</i>).....	31
3	Transformasi Data (<i>Data Transformation</i>)	33
4	Hasil Penelitian	35
	D. Kendala Penelitian.....	35
BAB IV Analisa Data		36
A.	K-Means Klasterisasi	36
B.	<i>Principal Component Analysis</i>	44
C.	Jaringan Syaraf Tiruan dengan Backpropagation.....	49
1	Implementasi Jaringan Syaraf Tiruan pada Klaster dengan Dimensi Tidak Tereduksi	49
2	Implementasi Jaringan Syaraf Tiruan pada Klaster dengan Dimensi Tereduksi.....	59
BAB V Penutup		63
A.	Kesimpulan.....	63
B.	Saran	64
DAFTAR PUSTAKA		65
APPENDIX 1		71
APPENDIX 2		75

Daftar Tabel

Tabel 3.1 : Daftar Perusahaan LQ – 45 Periode Agustus 2012 – Januari 2013 ...	27
Tabel 3.2 : Jumlah perusahaan per Sektor	29
Tabel 3.3 : Banyaknya emiten tiap tahun yang tergabung dalam indeks LQ-45 .	29
Tabel 4.1 : Final Klaster Centers	39
Tabel 4.2 : Jumlah Data Tiap Klaster	40
Tabel 4.3 : Matrik Komponen Tiap Klaster	46
Tabel 4.4 : Contoh Masukan untuk data Pelatihan Klaster 1	49
Tabel 4.5 : Contoh Keluaran untuk data Pelatihan Klaster 1.....	50
Tabel 4.6: Matriks Bobot Layer Masukan ke Layer 1	55
Tabel 4.7: Matriks Bobot Layer 1 ke Layer 2	56
Tabel 4.8: Matriks Bobot Layer 2 ke Layer 3 (keluaran)	56
Tabel 4.9: Hasil keluaran berbanding hasil keluaran jaringan klaster 1 – 3	57
Tabel 4.10: Hasil kinerja jaringan percobaan dengan data pengujian	58
Tabel 4.11 : Hasil Pengujian terhadap jaringan klaster 4	58
Tabel 4.12: Jaringan Syaraf Baru dengan hasil Principal Component Analysis .	60
Tabel 4.13: Matriks Bobot Klaster 1 (Layer Masukan ke Layer 1)	60
Tabel 4.14 : Matriks Bobot Klaster 5 (Layer 1 ke Layer Keluaran)	60
Tabel 4.15 : Matriks Bobot Klaster 5 (Layer Masukan ke Layer 1)	60
Tabel 4.16 : Matriks Bobot Klaster 5 (Layer 1 ke Layer 2)	60
Tabel 4.17 : Matriks Bobot Klaster 5 (Layer 2 ke Layer Keluaran)	60
Tabel 4.18: Perbandingan MSE (Kinerja) antara jaringan syaraf dengan PCA dan tanpa PCA	62

Daftar Gambar

Gambar 2.1: Diagram Alir Metode K-Means	20
Gambar 2.2 : Arsitektur Backpropagation	24
Gambar 3.1 : Contoh Script SQL Pembersihan Data	32
Gambar 3.2: Contoh Script SQL transformasi data	34
Gambar 4.1: Halaman Awal tampilan Database Wizards	37
Gambar 4.2 : Tampilan Awal Dataset pada SPSS	38
Gambar 4.3: Metode Principal Component Analysis pada SPSS	45
Gambar 4.4: Metode Principal Component Analysis pada SPSS	45
Gambar 4.5 : Hasil Performace Rate Pada Klaster 1	52
Gambar 4.6: Training State Klaster 1	53
Gambar 4.7: Analisa Regresi pada Jaringan Syaraf Tiruan untuk Klaster 1	53
Gambar 4.8: Analisa Regresi pada Jaringan Syaraf Tiruan untuk Klaster 4	54
Gambar 4.9 : Analisa Regresi pada Jaringan Syaraf Tiruan untuk Klaster 7.....	55