

**ANALISIS GEROMBOL DERET WAKTU DALAM
PENENTUAN MODEL PERAMALAN**

**Studi Kasus: 15 Saham (dalam negeri) Terindeks LQ45 dan
Sektoral Industri**

Tugas Tengah Semester
Kapita Selekta Statistika Terapan
(STK 691)

**RIZQI HARYASTUTI
G152180251**



**SEKOLAH PASCASARJANA
INSTITUT PERTANIAN BOGOR
BOGOR
2019**

PENDAHULUAN

Latar Belakang

Investasi atau penanaman modal merupakan suatu istilah yang berkaitan dengan akumulasi suatu bentuk aktiva dengan suatu harapan mendapatkan keuntungan pada masa depan. Berdasarkan teori ekonomi, investasi berarti pembelian dari modal barang yang tidak dikonsumsi tetapi digunakan untuk produksi yang akan datang. Ada berbagai bentuk investasi, seperti investasi tanah, pendidikan, dan saham. Salah satu investasi yang cukup hangat diperbincangkan adalah investasi saham, yakni kegiatan penanaman modal berupa pembelian saham atau surat berharga perusahaan. Bursa Efek Indonesia (BEI) adalah tempat bagi para pelaku saham untuk bertransaksi, baik menjual ataupun membeli saham, yakni salah satu produk pasar modal dan instrumen investasi jangka panjang. Perusahaan yang sahamnya dibeli oleh investor telah tercatat oleh BEI dan biasa disebut juga emiten, dimana setiap emiten punya kode saham atau simbol masing-masing.

Dalam melakukan kegiatan investasi tentunya kenaikan dan penurunan harga saham dari hari ke hari menjadi suatu hal yang penting diperhatikan karena diperlukan untuk pengambilan keputusan antara membeli atau menjual. Perubahan harga ini terus berubah tergantung pada penawaran dan permintaan saham. Pada beberapa emiten ditemui pergerakan harga saham sangat ekstrim dan nonlinear, sehingga meramalkan harga dan menentukan keputusan waktu untuk membeli atau menjual menjadi hal yang tidak mudah. Mengingat beberapa risiko investasi saham, antara lain adanya *capital loss* atau kerugian atas penurunan harga saham, risiko likuidasi apabila perusahaan bangkrut, dan tidak mendapatkan dividen, membuat sebagian besar orang awam takut melakukan investasi ini. Tidak jarang ditemukan beberapa emiten memiliki pola historis saham yang sangat berbeda satu sama lain. Pada penelitian ini akan dipelajari pola-pola dari data historis saham dan dilakukan penggerombolan terhadap pola-pola historis yang mirip.

Data deret waktu adalah tipe umum dari data dinamis yang muncul secara alami dalam banyak skenario, salah satunya pada data saham. Deret waktu menimbulkan tantangan tersendiri jika periodenya panjang dan data memiliki dimensi yang tinggi. Analisis gerombol deret waktu adalah salah satu jenis algoritma penggerombolan yang dibuat untuk menangani data dinamis. Aghabozorgi *et al.* (2015) menyatakan elemen yang paling penting dipertimbangkan adalah ukuran jarak ketidaksamaan, fungsi ekstraksi dasar (jika ada), algoritma penggerombolan itu sendiri, dan evaluasi gerombol.

Sebagai pemula dalam melakukan investasi, banyaknya pilihan saham tidak jarang membingungkan. Terdapat berbagai saham dengan beragam fluktuasi historis. Pada penelitian ini akan dilakukan analisis penggerombolan deret waktu terhadap beberapa harga saham milik emiten ternama yang tercantum sebagai bagian dari indeks LQ45 sekaligus sektoral industri. Diduga saham yang berasal dari emiten dalam sektoral industri yang sama memiliki pasang surut historis yang sama, sehingga dalam melakukan peramalan ke depannya saham yang berasal dari emiten dalam sektoral industri yang sama tersebut dapat menjadi satu anggota gerombol yang sama dan menggunakan model peramalan yang sama. Hal ini tentu lebih efektif dibandingkan dengan peramalan secara individu masing-masing saham.

Tujuan Penelitian

Tujuan dilakukannya penelitian ini adalah untuk mengetahui ukuran penggerombolan deret waktu yang lebih baik saat menggunakan teknik hirarki dan dengan metode pautan; membandingkan efektifitas dan kebaikan peramalan yang dilakukan secara individu dan gerombol.

TINJAUAN PUSTAKA

Daftar Istilah

Indeks Saham LQ45

Indeks LQ45 adalah indeks yang mengukur performa harga 45 saham yang memiliki likuiditas tinggi dan kapitalisasi pasar besar serta didukung oleh fundamental perusahaan yang baik.

Indeks Saham Sektoral

Indeks sektoral adalah indeks yang mengukur performa harga seluruh saham dari masing-masing sektor industri yang terdapat pada klasifikasi *Jakarta Stock Industrial Classification* (JASICA) yang terdiri dari:

- Indeks pertanian
- Indeks pertambangan
- Indeks industri dasar dan kimia
- Indeks aneka industri
- Indeks industri barang konsumsi
- Indeks properti, *real estate*, dan konstruksi bangunan
- Indeks infrastruktur, utilitas, dan transportasi
- Indeks keuangan
- Indeks perdagangan, jasa, dan investasi
- Indeks manufaktur

Closing Price

Kenton (2018) menyatakan *closing price* atau harga penutupan adalah harga terakhir atau final dimana sekuritas diperdagangkan pada setiap hari perdagangan tertentu. *Closing price* menggambarkan penilaian terkini dari sebuah sekuritas hingga perdagangan dimulai kembali pada hari berikutnya. Sebagian besar instrumen keuangan diperdagangkan setelah jam kerja meskipun dengan volume dan tingkat likuiditas yang sangat kecil. Namun dikarenakan *closing price* pada suatu sekuritas mungkin tidak sesuai dengan harga setelah kerjanya, muncul *adjusted closing price* yang dapat menggambarkan efek aksi perusahaan terhadap harga saham.

Adjusted Closing Price

Kenton (2018) menyatakan *adjusted closing price* atau harga penutupan yang disesuaikan adalah harga penutupan saham pada setiap hari perdagangan yang telah diamandemen untuk memasukkan setiap distribusi dan aksi dari perusahaan atau korporasi yang terjadi sebelum pembukaan harga pada hari berikutnya. *Adjusted closing price* sering digunakan ketika memeriksa *return* historis atau melakukan analisis yang rinci terhadap *return* historis.

Analisis Gerombol Deret Waktu

Clustering atau penggerombolan merupakan bagian dari *unsupervised learning* yang bertujuan untuk membagi sejumlah objek yang tidak memiliki label atau kelas menjadi kelompok-kelompok yang homogen atau *clusters*. Penggerombolan dilakukan sedemikian rupa sehingga objek-objek dalam kelompok yang sama adalah lebih mirip satu sama lain dibandingkan objek lain yang berada pada kelompok yang berbeda berdasarkan kriteria-kriteria tertentu. Komponen utama dalam analisis gerombol adalah menentukan ukuran ketidaksamaan atau *dissimilarities* yang tepat antara dua objek data (Montero dan Vilar 2014).

Pendekatan yang digunakan untuk mendefinisikan ukuran ketidaksamaan dalam analisis gerombol pada data yang tidak mengandung unsur waktu dan data deret waktu tentu berbeda. Hal ini dikarenakan karakter data deret waktu yang dinamis dan pada penggerombolan konvensional hubungan saling ketergantungan antar nilai-nilai atau *autocorrelation* akan diabaikan. Menurut Aghabozorgi *et al.* (2014) metode algoritma yang telah dikembangkan untuk analisis gerombol deret waktu pada dasarnya mempunyai tiga karakteristik utama. Pertama dengan memodifikasi algoritma untuk data *cross section* agar dapat digunakan pada data deret waktu. Kedua dengan mengonversi data deret waktu menjadi *cross section* sehingga algoritma analisis gerombol data *cross section* dapat diterapkan. Ketiga menggunakan pendekatan *multi-step*, yaitu kombinasi dari kedua pendekatan sebelumnya. Hal tersebut analog dengan empat kategori pengukuran ketidaksamaan yang dinyatakan oleh Montero dan Vilar (2014), yakni pendekatan bebas model yang mempertimbangkan matriks konvensional berdasarkan kedekatan nilainya pada titik waktu tertentu, pendekatan berdasar model yang mengasumsikan model dibentuk dengan struktur parametrik tertentu, pendekatan berbasis kompleksitas yakni pengukuran ketidaksamaan berdasarkan perbandingan tingkat kerumitan serial, pendekatan berbasis prediksi.

Penggerombolan Hirarki

Penggerombolan hirarki adalah sebuah algoritma yang berusaha membentuk kelompok-kelompok hirarki yang dimana seiring peningkatan level hirarki maka kelompok-kelompok akan terbentuk dengan menggabungkan kelompok dari tingkat bawah selanjutnya sehingga urutan-urutan pengelompokan diperoleh. Untuk menentukan bagaimana penggerombolan dilakukan, ukuran ketidaksamaan antar kelompok harus ditentukan. Namun jumlah gerombol tertentu tidak perlu ditentukan terlebih dahulu karena algoritma hirarki bersifat aglomerasi atau memecah belah. Dalam prosedur aglomerasi setiap anggota data mulai dalam gerombol sendiri dan anggota dikelompokkan bersama secara berurutan berdasarkan ukuran kesamaan sampai semua anggota terkandung dalam satu gerombol. Hirarki yang dibuat divisualisasikan sebagai pohon biner dimana ketinggian setiap simpul sebanding dengan nilai perbedaan antarkelompok antara dua anak simpul, plot ini disebut dendrogram (Hastie *et al.* 2009). Dalam algoritma hirarki terdapat beberapa ukuran jarak yang dapat digunakan untuk membentuk gerombol, dalam tulisan ini akan dibahas beberapa ukuran yang erat kaitannya dengan analisis gerombol deret waktu dan penerapannya menggunakan R serta termasuk pendekatan bebas model (Montero dan Vilar 2014), yakni jarak berbasis korelasi (*correlation-based distances*), jarak Frechet, dan jarak perbelokan waktu yang dinamis (*dynamic time warping distance*).

Frechet Distance

Jarak ini diperkenalkan oleh Fréchet (1906) untuk mengukur kedekatan antara kurva kontinu, tetapi telah banyak digunakan pada kasus diskrit (lihat Eiter dan Mannila 1994) dan dalam kerangka deret waktu. Definisi formal untuk kasus diskrit dapat diberikan sebagai berikut. Misalkan M adalah himpunan dari semua urutan yang mungkin dari pasangan m yang menjaga urutan pengamatan dalam bentuk

$$r = ((X_{a_1}, Y_{b_1}), (X_{a_2}, Y_{b_2}), \dots, (X_{a_m}, Y_{b_m}))$$

$$d_F(\mathbf{X}_T, \mathbf{Y}_T) = \min_{r \in M} \left(\max_{i=1, \dots, m} |X_{a_i} - Y_{b_i}| \right)$$

Dynamic Time Warping Distance

Dynamic Time Warping (DTW) distance merupakan ukuran jarak berdasarkan pembelokan waktu yang dinamis untuk menemukan pemetaan r antar waktu sehingga meminimumkan ukuran jarak antar observasi (X_t, Y_t) , dimana r pada DTW sama dengan *Frechet*. Espinosa (2017) menyatakan bahwa DTW adalah algoritma pemrograman dinamis yang membandingkan dua seri dan mencoba untuk menemukan jalur pembelokan yang optimal di antara mereka.

$$d_{DTW}(\mathbf{X}_T, \mathbf{Y}_T) = \min_{r \in M} \left(\sum_{i=1, \dots, m} |X_{a_i} - Y_{b_i}| \right)$$

Correlation-based Distance

Correlation-based distance merupakan ukuran kriteria ketidaksamaan sekaligus kesamaan yang pertama dikembangkan dan sederhana dengan mempertimbangkan faktor korelasi Pearson \bar{X}_T dan \bar{Y}_T , dimana \bar{X}_T dan \bar{Y}_T merupakan nilai rata-rata dari realisasi seri \mathbf{X}_T dan \mathbf{Y}_T itu sendiri.

$$COR(\mathbf{X}_T, \mathbf{Y}_T) = \frac{\sum_{t=1}^T (\mathbf{X}_T - \bar{X}_T)(\mathbf{Y}_T - \bar{Y}_T)}{\sqrt{\sum_{t=1}^T (\mathbf{X}_T - \bar{X}_T)^2} \sqrt{\sum_{t=1}^T (\mathbf{Y}_T - \bar{Y}_T)^2}}$$

$$d_{COR}(\mathbf{X}_T, \mathbf{Y}_T) = \sqrt{2(1 - COR(\mathbf{X}_T, \mathbf{Y}_T))}$$

Model Autoregressive Integrated Moving Average (ARIMA)

Suatu deret waktu $\{Y_t\}$ adalah suatu model campuran *Autoregressive Integrated Moving Average* apabila pada pembedaan ke- d merupakan proses ARMA yang stasioner. Jadi, jika $\{W_t\}$ mengikuti model ARMA (p, q) , maka $\{Y_t\}$ merupakan suatu proses ARIMA (p, d, q) (Cryer dan Chan 2008). Notasi umum ARIMA yaitu:

$$\text{ARIMA}(p, d, q)$$

dengan:

p, q : order AR, MA yang dari model
 d : order *differencing*

Bentuk matematis dari model ARIMA (p, d, q) yaitu sebagai berikut:

$$\varphi_p(B)(1 - B)^d = \theta_q(B) e_t$$

dengan :

$\varphi_p(B)$: komponen AR non-musiman

$\theta_q(B)$: komponen MA non-musiman
 $(1 - B)^d$: operator proses pembedaan
 e_t : error
 B : operator *Backshift*

Evaluasi Peramalan

Peramalan untuk beberapa periode ke depan dilakukan dengan menggunakan model yang telah dibuat. Pengguna mengharapkan data aktual dan peramalan mempunyai beda sekecil mungkin. Semakin kecil perbedaannya, maka dikatakan semakin akurat. Salah satu ukuran evaluasi model peramalan yang dapat digunakan adalah *Mean Absolute Percentage Error* (MAPE). Pengukuran akurasi ini kerap dipilih karena terbebas dari skala yang digunakan pada amatan asli. Definisi MAPE dalam Montgomery *et al.* (2008) adalah:

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{y_t - \hat{y}_t}{y_t} \right|$$

METODOLOGI PENELITIAN

Data

Dalam penelitian ini digunakan 15 data saham dalam negeri yang tercatat secara harian pada 1 Agustus 2018 hingga 31 Desember 2018. Data ini merupakan data sekunder yang diperoleh secara *online* melalui situs <https://finance.yahoo.com/>. Perusahaan emiten yang digunakan datanya pada penelitian ini adalah perusahaan yang masuk dalam Indeks Sektoral sekaligus Indeks LQ45. Tidak semua data saham dalam situs digunakan, melainkan hanya *adjusted close price*. Berikut adalah daftar nama saham yang digunakan:

Tabel 1. Daftar saham dalam negeri yang digunakan

Simbol	Nama Perusahaan	Sektor
ACES	PT Ace Hardware Indonesia Tbk	Perdagangan, Jasa, dan Investasi
MNCN	PT Media Nusantara Citra Tbk	
ANTM	PT Aneka Tambang (Persero) Tbk	Pertambangan
ELSA	PT Elnusa Tbk	
PTBA	PT Tambang Batubara Bukit Asam Tbk	
PGAS	Perusahaan Gas Negara (Persero) Tbk	Infrastruktur, Utilitas, dan Transportasi
JSMR	Jasa Marga (Persero) Tbk	
EXCL	XL Axiata Tbk	
TLKM	Telekomunikasi Indonesia Tbk	
GIAA	Garuda Indonesia (Persero) Tbk	
INDF	PT Indofood Sukses Makmur Tbk	Barang konsumsi
KAEF	PT Kimia Farma (Persero) Tbk	
KLBF	PT Kalbe Farma Tbk	
ADHI	PT Adhi Kaya (Persero) Tbk	Properti, Real Estate, dan Konstruksi Bangunan
ASRI	PT Alam Sutera Realty Tbk	

Prosedur Analisis Data

Analisis data dalam penelitian ini menggunakan perangkat lunak statistika R 3.4.0 dan RStudio. Adapun *package* yang diperlukan: *quantmod* untuk mengunduh data saham dari situs Yahoo secara langsung, *xts* untuk menyimpan data saham yang merupakan data panel, dan *Tsclust* untuk melakukan analisis gerombol pada data saham. Selain itu dibutuhkan paket *zoo*, *ggplot*, *reshape2*, *tseries*, *forecast*, *TSA*, dan *TTR*. Berikut adalah tahapan analisis data:

1. Mengunduh data saham yaitu *adjusted close price* dari situs Yahoo menggunakan *quantmod* dengan rincian sebagai berikut:
 - i. Inisialisasi simbol saham yang akan diunduh.
 - ii. Deklarasi rentang waktu saham, yakni mulai 1 Agustus hingga 31 Desember 2018.

2. Menyimpan data yang telah diunduh dalam format *xts*.
3. Melakukan inisialisasi terhadap data saham yang hilang (bernilai NA) dengan rata-rata dari saham itu sendiri.
4. Melakukan eksplorasi secara visual dilakukan dengan plot deret waktu dan diagram kotak garis dari setiap saham dalam negeri dalam periode 1 Agustus hingga 31 Desember 2018 untuk melihat gambaran umum pergerakan setiap saham dari waktu ke waktu.
5. Melakukan transformasi logaritma natural terhadap setiap data saham agar memiliki sebaran yang sama dan homogen.
6. Membagi data saham menjadi 2, yakni data latih (*training data*) untuk membentuk model dan penggerombolan dan data uji (*testing data*) untuk validasi, dimana data latih memiliki periode 1 Agustus hingga 31 Desember 2018, sedangkan data uji pada 1 hingga 31 Januari 2019.
7. Pemodelan ARIMA terhadap saham secara individu.
 - i. Mengecek kestasioneran masing-masing saham dengan menggunakan *Dicky Fuller Test*.
 - ii. Melakukan *differencing* pada data saham yang tidak stasioner.
 - iii. Identifikasi model tentatif ARIMA (p,d,q) berdasarkan plot ACF dan PACF, serta EACF.
 - iv. Memilih model terbaik untuk tiap saham berdasarkan AIC terkecil.
8. Peramalan data latih dengan model individu untuk validasi model terhadap data tes.
 - i. Melakukan peramalan pada masing-masing saham dalam data latih berdasarkan model individu untuk periode 1 hingga 31 Januari 2019 (sesuai periode pada data tes).
 - ii. Menghitung MAPE masing-masing saham, dimana persentase relatif merupakan perbandingan absolut antara *forecast error* dengan data aktual (data tes).
9. Penggerombolan terhadap 15 data saham dalam negeri dengan 3 jarak yang berbeda, yaitu *Correlation*, *Frechet*, dan *Dynamic Time Warping*.
 - i. Mengubah format kumpulan saham yang semula berupa *.xts* menjadi matriks.
 - ii. Menghitung jarak kemiripan antar saham menggunakan fungsi *diss()* pada paket *Tsclust* dengan metode *COR* yang berarti *correlation*.
 - iii. Melakukan penggerombolan hirarki berdasarkan jarak tersebut menggunakan fungsi *hclust()* dengan metode pautan rataan.
 - iv. Membuat plot dendrogram yang memudahkan identifikasi anggota-anggota gerombol.
 - v. Selanjutnya kembali dilakukan perhitungan jarak, namun kali ini jarak ketidakmiripan antar saham dengan metode *FRECHET* pada fungsi dan paket yang sama dengan 8.ii.
 - vi. Melakukan hal yang sama seperti 8.iii. dan 8.iv. untuk metode *FRECHET*.
 - vii. Perhitungan jarak ketidakmiripan antar saham dengan metode *DTWARP* atau *Dynamic Time Warping* pada fungsi dan paket yang sama dengan 8.ii.
 - viii. Melakukan hal yang sama seperti 8.iii. dan 8.iv. untuk metode *DTWARP*.
 - ix. Untuk setiap metode perhitungan jarak, diambil *k* atau jumlah gerombol dari dendrogram sebesar 5 untuk menyesuaikan banyaknya sektor industri saham yang digunakan pada penelitian ini.

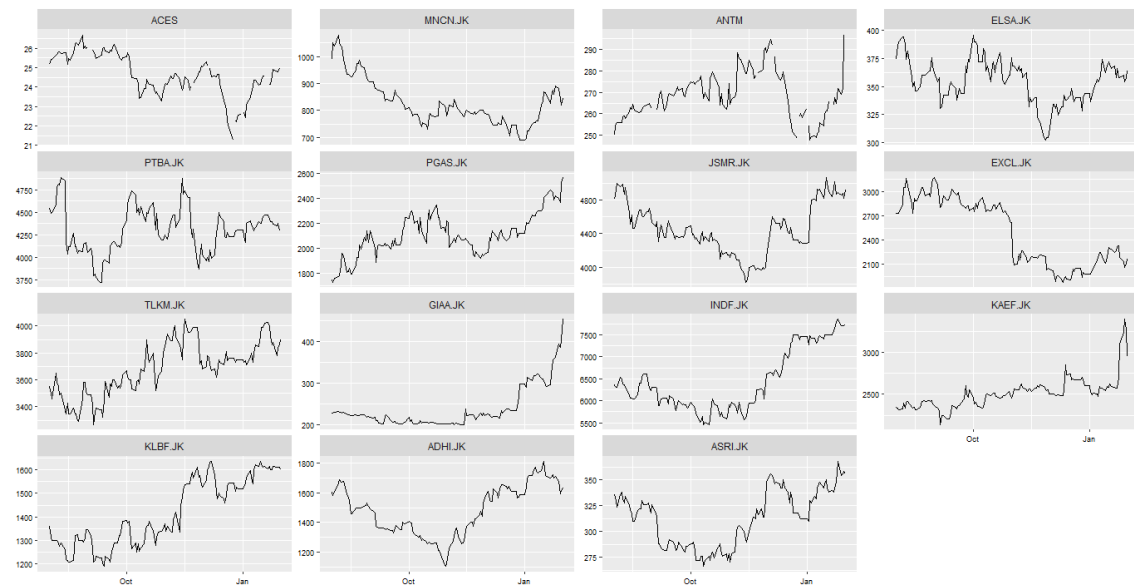
10. Pemodelan ARIMA secara gerombol. Karena metode yang digunakan dalam penggerombolan adalah pautan rata-rata, maka digunakan rata-rata dalam menentukan model perwakilan tiap gerombol.
 - i. Menghitung rata-rata saham untuk setiap indeks waktu ke- t sesuai dengan keanggotaan pada gerombol ke- k , dimana $k = 1, 2, \dots, 5$, yang dihasilkan metode *Correlation*. Karena data yang digunakan adalah data latih, maka rata-rata saham pada setiap gerombol berjalan dari 1 Agustus hingga 31 Desember 2018.
 - ii. Membuat plot deret waktu untuk tiap gerombol.
 - iii. Analog dengan pemodelan saham secara individu, langkah 7.i. hingga 7.iv. juga dilakukan terhadap masing-masing gerombol data saham.
 - iv. Selanjutnya melakukan langkah 9.i. hingga 9.iii. berturut-turut pada gerombol-gerombol yang dihasilkan metode *Frechet* dan *Dynamic Time Warping*.
 - v. Menghitung rata-rata AIC dari 5 gerombol saham untuk setiap metode jarak penggerombolan. Penggerombolan yang memiliki rata-rata AIC model lebih kecil yang akan digunakan untuk peramalan selanjutnya.
11. Peramalan data latih dengan model gerombol untuk validasi model terhadap data tes. Setelah didapatkan ukuran jarak penggerombolan yang sesuai, model ARIMA gerombol digunakan untuk peramalan saham secara individu sesuai keanggotaannya pada gerombol tersebut. Misalkan saham ACES adalah anggota gerombol ke-3, maka saham ACES diramalkan dengan model ARIMA gerombol ke-3. Proses peramalan secara individu analog dengan langkah 8.i. dan 8.ii., namun model yang digunakan adalah model gerombol.
12. Bandingkan MAPE dan rata-rata MAPE yang dihasilkan dari penggunaan model secara individu dan secara gerombol pada masing-masing saham. Nilai MAPE dan rata-rata MAPE yang lebih kecil menunjukkan pendekatan yang lebih baik untuk peramalan dengan pemodelan pada 15 saham dalam negeri, apakah dengan model secara individu atau gerombol.

HASIL DAN PEMBAHASAN

Eksplorasi Data

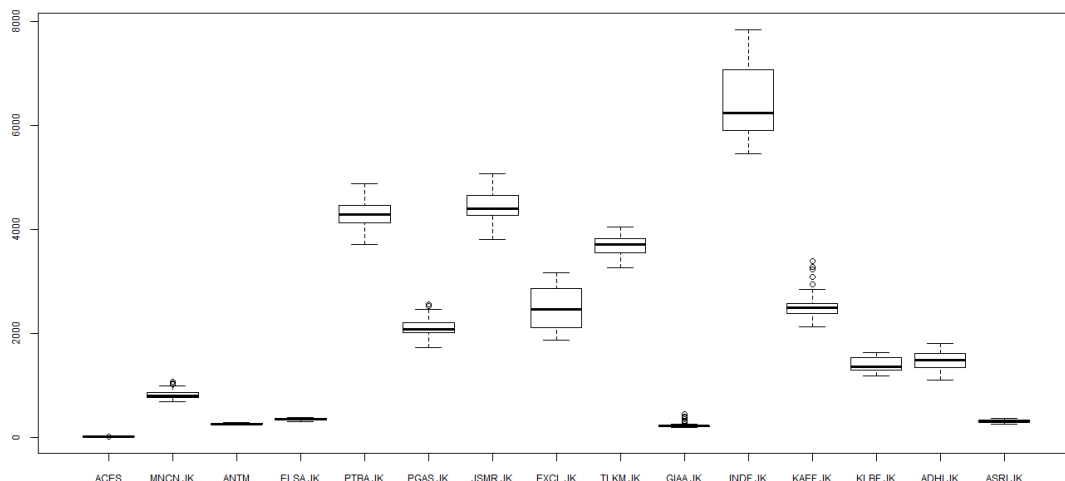
Sebelum melakukan analisis lebih jauh, dilakukan eksplorasi terhadap 15 saham dalam negeri untuk mendapatkan gambaran umum mengenai pergerakan dari waktu ke waktu. Berikut adalah plot deret waktu masing-masing saham pada periode 1 Agustus 2018 hingga 31 Januari 2019 dan diagram kotak garis:

Gambar 1. Plot deret waktu masing-masing saham



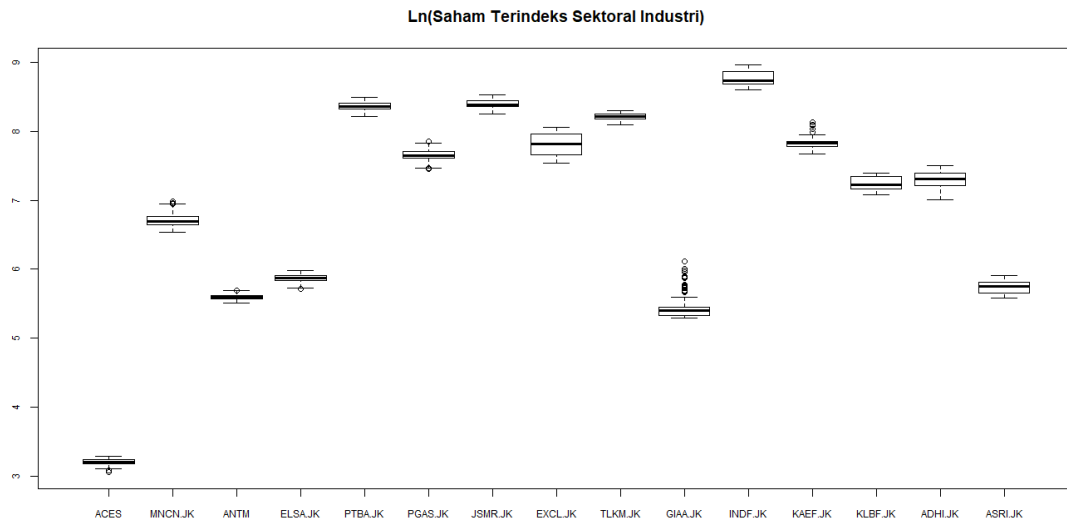
Gambar 2. Diagram kotak garis masing-masing saham

Saham Terindeks Sektoral Industri



Berdasarkan kedua gambar tersebut, dapat dilihat bahwa masing-masing saham memiliki ragam yang besar. Sehingga untuk memperkecil besarnya ragam tersebut dilakukan transformasi logaritma natural atau 'ln' terhadap setiap saham. Berikut adalah tampilan diagram kotak garis masing-masing saham yang telah ditransformasi:

Gambar 3. Diagram kotak garis masing-masing $\ln(\text{saham})$



Berdasarkan gambar tersebut telah didapatkan ragam setiap saham yang lebih sempit dibanding data aktual. Namun ragam yang mengecil belum menjamin stasioneritas data, seperti pada 15 saham ini (plot deret waktu setelah transformasi \ln tidak ditampilkan).

Tabel 2. Ringkasan stasioneritas masing-masing $\ln(\text{saham})$

Simbol Saham	p-value	Kesimpulan
ACES	0.4497	Data tidak stasioner
MNCN	0.3726	Data tidak stasioner
ANTM	0.4993	Data tidak stasioner
ELSA	0.1866	Data tidak stasioner
PTBA	0.0851	Data tidak stasioner
PGAS	0.2062	Data tidak stasioner
JSMR	0.3942	Data tidak stasioner
EXCL	0.4205	Data tidak stasioner
TLKM	0.2708	Data tidak stasioner
GIAA	0.9900	Data tidak stasioner
INDF	0.9570	Data tidak stasioner
KAEF	0.0100	Data stasioner
KLBF	0.1938	Data tidak stasioner
ADHI	0.6492	Data tidak stasioner
ASRI	0.5500	Data tidak stasioner

Sehingga untuk mengatisinya dilakukan pengecekan ulang dengan `adf.test()` secara individu dan *differencing* bagi data saham yang tidak stasioner.

Pemodelan ARIMA secara Individu

Untuk mendapatkan model tentatif yang tepat bagi setiap saham, tentu dilakukan pengecekan terhadap beberapa kandidat model ARIMA (p,d,q) yang diidentifikasi melalui plot ACF dan PACF serta EACF dari masing-masing saham yang telah stasioner. Berikut adalah ringkasan model ARIMA (p,d,q) yang telah tepat dari beberapa kandidat model untuk masing-masing saham dalam data latih:

Tabel 3. Ringkasan model ARIMA secara individu

Simbol Saham	ARIMA			AIC
	p	d	q	
ACES	0	1	1	-532.32
MNCN	4	1	0	-521.33
ANTM	0	1	1	-541.84
ELSA	1	1	1	-479.63
PTBA	1	1	1	-452.07
PGAS	4	1	0	-457.16
JSMR	1	1	0	-546.3
EXCL	0	1	5	-473.83
TLKM	0	1	2	-541.67
GIAA	1	1	1	-442.75
INDF	1	1	1	-528.17
KAEF	1	0	0	-524.74
KLBF	1	1	1	-499.44
ADHI	0	1	1	-515.18
ASRI	1	1	1	-512.32

Model yang telah didapatkan untuk masing-masing saham tersebut nantinya akan digunakan untuk melakukan peramalan lalu validasi dengan data tes.

Penggerombolan Data Deret Waktu

Pada penelitian ini digunakan penggerombolan menggunakan teknik hirarki dengan 3 jenis jarak atau ukuran, yaitu *Correlation*, *Frechet*, dan *Dynamic Time Warping* sebagai perbandingan untuk mengetahui ukuran yang paling sesuai dalam menggambarkan saham-saham.

Tabel 4 adalah ringkasan anggota gerombol dari plot dendogram penggerombolan menggunakan 3 jarak yang berbeda. Berdasarkan ringkasan tersebut dapat diketahui bahwa hasil penggerombolan dengan jarak *Frechet* dan *Dynamic Time Warping* adalah sama. Selanjutnya digunakan *Dynamic Time Warping* saja untuk dibandingkan dengan *Correlation*. Perbandingan dilakukan dengan melakukan peramalan data latih secara gerombol kemudian dilakukan validasi terhadap data tes. Ukuran yang sesuai digunakan dalam penggerombolan 15 saham akan menghasilkan AIC yang lebih kecil.

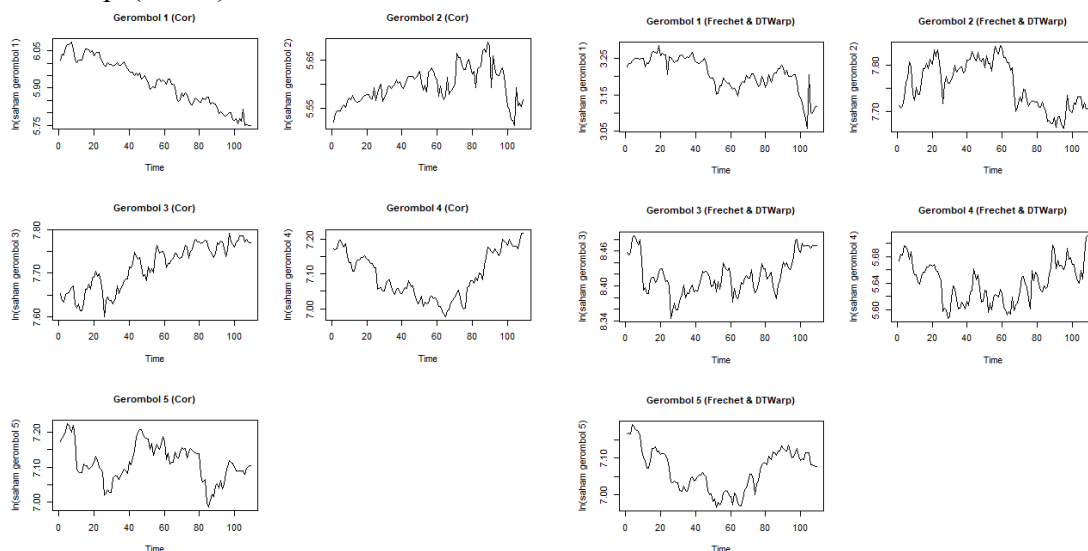
Tabel 4. Ringkasan hasil penggerombolan

Simbol Saham	Gerombol		
	<i>Correlation</i>	<i>Frechet</i>	<i>DTWarp</i>
ACES	1	1	1
MNCN	1	5	5
ANTM	2	4	4
ELSA	5	4	4
PTBA	5	3	3
PGAS	3	2	2
JSMR	4	3	3
EXCL	1	2	2
TLKM	3	3	3
GIAA	4	4	4
INDF	4	3	3
KAEF	3	2	2
KLBF	3	5	5
ADHI	4	5	5
ASRI	4	4	4

Pemodelan ARIMA secara Gerombol

Dalam melakukan pemodelan ARIMA secara gerombol, terlebih dahulu kita menentukan perwakilan gerombol yang akan dimodelkan. Pada penelitian ini digunakan pautan rata-rata sebagai perwakilan dari anggota-anggota di setiap gerombol pada periode 1 Agustus hingga 31 Desember 2018 (data latih). Berikut adalah plot deret waktu untuk setiap gerombol, baik yang menggunakan *Correlation* dan *DTWarp*:

Gambar 4. Plot deret waktu gerombol saham dengan jarak *Correlation* (kiri) dan *DTWarp* (kanan)



Berdasarkan plot masing-masing gerombol dari dua metode jarak penggerombolan yang dibandingkan, semua plot deret waktu menunjukkan kecenderungan gerombol saham yang tidak stasioner. Sehingga sebelum melakukan peramalan untuk periode 1 hingga 31 Januari 2019, dilakukan *differencing* terlebih dahulu. Selanjutnya untuk proses pemodelan secara gerombol, langkah yang dilakukan analog dengan pemodelan secara individu. Berikut adalah ringkasan model secara gerombol:

Tabel 5. Ringkasan pemodelan secara gerombol

Jarak	Gerombol	ARIMA			AIC	Rata-rata AIC
		p	d	q		
<i>Correlation</i>	1	1	0	0	-579.31	-563.742
	2	0	1	1	-541.84	
	3	1	1	0	-584.94	
	4	1	1	1	-584.68	
	5	0	1	5	-527.94	
<i>DTWarp</i>	1	0	1	1	-532.32	-575.314
	2	5	1	0	-544.11	
	3	1	0	0	-610.02	
	4	1	1	0	-595.14	
	5	1	1	1	-594.98	

Tabel 6. Rincian model ARIMA secara gerombol

Simbol Saham	Gerombol	ARIMA		
		p	d	q
ACES	1	0	1	1
MNCN	5	1	1	1
ANTM	4	1	1	0
ELSA	4	1	1	0
PTBA	3	1	0	0
PGAS	2	5	1	0
JSMR	3	1	0	0
EXCL	2	5	1	0
TLKM	3	1	0	0
GIAA	4	1	1	0
INDF	3	1	0	0
KAEF	2	5	1	0
KLBF	5	1	1	1
ADHI	5	1	1	1
ASRI	4	1	1	0

Dapat dilihat dari ringkasan model ARIMA (p,d,q) masing-masing gerombol berdasarkan jarak yang digunakan bahwa model pada anggota gerombol berukuran

DTWarp menghasilkan rata-rata AIC yang lebih kecil, yang berarti ukuran *DTWarp* dianggap lebih mampu menggambarkan 15 saham yang digerombolkan. Selanjutnya tahap peramalan data latih secara individu untuk validasi terhadap data tes dilakukan dengan model-model pada 5 gerombol berukuran *DTWarp*. Tabel adalah rincian model sesuai keanggotaan saham pada gerombol *DTWarp*.

Validasi Model

Pada tahap validasi model, akan dilakukan peramalan data latih sepanjang periode yang sama dengan data tes, yakni 1 hingga 31 Januari 2019. Kemudian perbandingan nilai ramalan dan aktual tersebut, yaitu *Mean Absolute Percentage Error* (MAPE), menjadi ukuran kebaikan model. Nilai MAPE yang lebih kecil menunjukkan pendekatan peramalan saham yang lebih baik, secara individu atau gerombol dengan *DTWarp*. Berikut adalah ringkasan perhitungan MAPE antara kedua pendekatan:

Tabel 7. Ringkasan perhitungan MAPE

Simbol Saham	MAPE Peramalan	
	Individu	Gerombol
ACES	2.17413	0.46640
MNCN	2.08984	2.13777
ANTM	0.57446	0.57449
ELSA	0.74812	0.74812
PTBA	0.20986	0.21145
PGAS	1.37043	1.30931
JSMR	1.40882	1.40882
EXCL	1.21393	1.21449
TLKM	0.36259	0.36525
GIAA	1.79762	1.79762
INDF	0.18880	0.19029
KAEF	0.94945	0.74267
KLBF	0.63339	0.63534
ADHI	0.96133	0.95342
ASRI	1.55290	1.55290
Rata-rata	1.08238	0.95389

Jika dilihat pada ringkasan tersebut, tidak semua saham lebih baik diramalkan dengan pendekatan individu saja atau gerombol saja. Namun secara rata-rata, peramalan 15 saham dalam negeri tersebut lebih baik diramalkan dengan pendekatan gerombol yang ditunjukkan dengan rata-rata MAPE yang lebih kecil.

KESIMPULAN DAN SARAN

Kesimpulan

Berikut adalah beberapa kesimpulan yang dapat dirangkum berdasarkan dugaan awal pada latar belakang dan tujuan penelitian, yaitu:

- Anggota gerombol yang dihasilkan pada masing-masing gerombol belum tentu merupakan saham yang berasal dari emiten dalam sektoral industri yang sama dengan kata lain pasang surut data historis saham dalam masing-masing jenis sektoral industri dapat berbeda satu sama lain.
- Dari ketiga ukuran penggerombolan yang dibandingkan, yaitu *Correlation*, *Frechet*, dan *DTWarp*, didapatkan rata-rata model yang lebih baik saat menggunakan ukuran *DTWarp*. Hal ini ditunjukkan dengan rata-rata AIC dari 5 gerombol yang lebih kecil.
- Peramalan yang dilakukan dengan memanfaatkan model secara gerombol adalah lebih baik dibandingkan peramalan secara individu, yang dibuktikan dengan MAPE yang lebih kecil saat validasi model. Dan peramalan dengan model secara gerombol dirasa dapat lebih efektif, karena pemodelan tidak perlu dilakukan satu per satu.

Saran

Beberapa saran yang diusulkan untuk penelitian selanjutnya adalah sebagai berikut:

- Dilakukan perhitungan k optimum, yakni penentuan jumlah gerombol optimum yang dapat menggambarkan data.
- Dilakukan perbandingan antara penggerombolan menggunakan teknik hirarki dan non-hirarki, karena dalam beberapa pustaka disebutkan bahwa teknik non-hirarki lebih mudah dan efektif.
- Dilakukan penggerombolan pada data yang mengandung musiman dan non-musiman untuk melihat apakah jarak dapat menangkap perbedaan karakter data tersebut.

DAFTAR PUSTAKA

- Aghabozorgi S, Shirkhorshidi AS, Wah TY. 2015. Time-series clustering – A decade review. *Information System*. 53: 16-38. doi: 10.1016/j.is.2015.04.007.
- Ardalova R. 2017. Analisis Harga Cabai Merah Besar di Pasar Eceran Jakarta dengan Menggunakan Analisis Gerombol Deret Waktu [skripsi]. Bogor (ID): Institut Pertanian Bogor.
- Espinosa AS. 2017. Comparing Time-Series Clustering Algorithms in R Using the dtwclust Package.
- Montero P, Vilar JA. 2014. Tsclust: An R Package for Time Series Clustering. 62(1): 1-43. doi: 10.18637/jss.v062.i01
<https://www.idx.co.id/produk/indeks/> [diakses pada 28 Maret 2018 pukul 10.40]

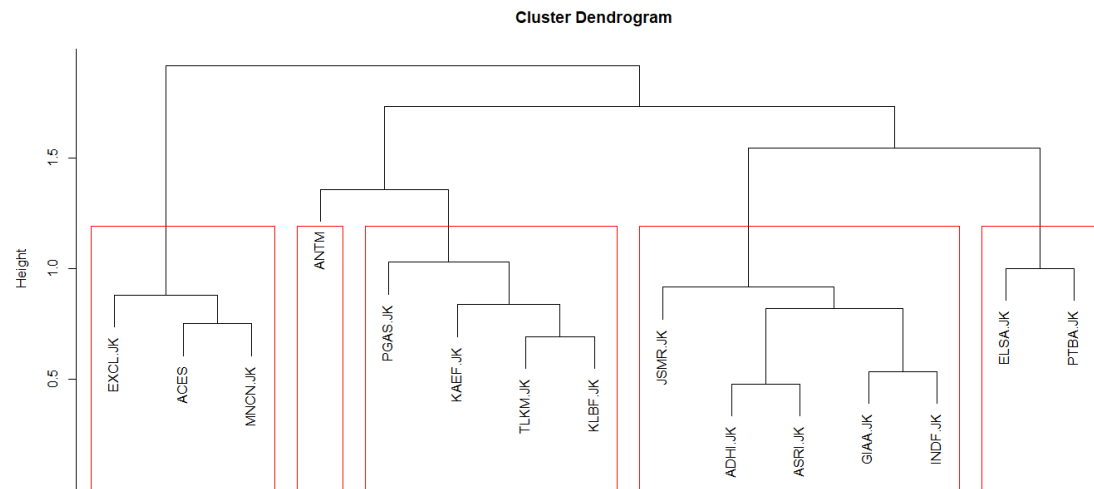
LAMPIRAN

Lampiran Data

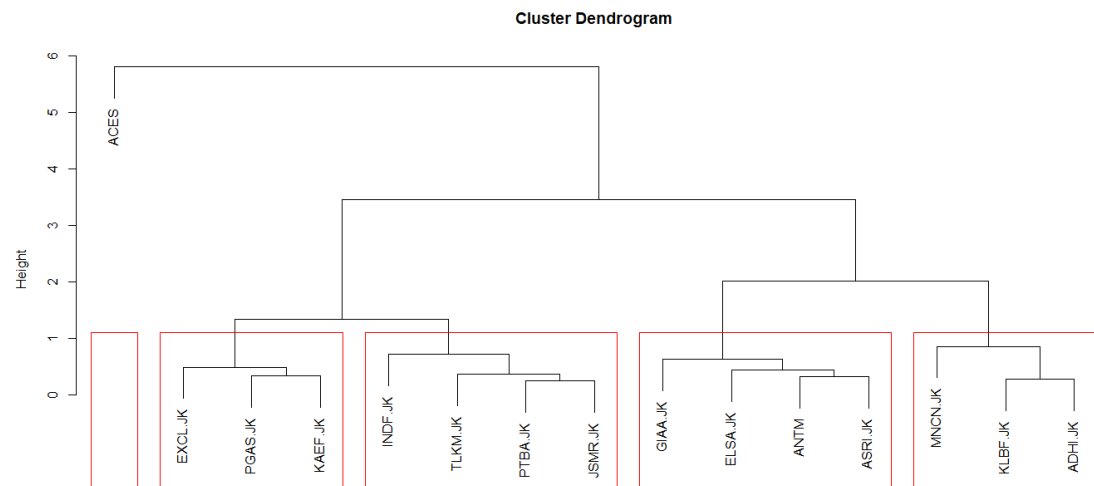
Time	ACES	MNCN.JK	ANTM	ELSA.JK	PTBA.JK	PGAS.JK	JSMR.JK	EXCL.JK	TLKM.JK	GIAA.JK	INDF.JK	KAEF.JK	KLBF.JK	ADHI.JK	ASRI.JK
8/1/2018	25.181101	990	250.27745	374	4540	1750	4810	2730	3550	228	6355.1045	2340	1360	1605	336
8/2/2018	25.444054	1050	254.5038	386	4490	1730	4860	2730	3500	228	6330.3765	2330	1315	1580	334
8/3/2018	25.451992	1040	255.91258	390	4500	1760	5000	2730	3460	230	6305.6484	2310	1300	1605	324
8/6/2018	25.635563	1080	256.09119	394	4580	1775	4960	2850	3650	232	6528.2007	2320	1300	1665	336
8/7/2018	25.709986	1050	255.93242	394	4780	1780	4980	3050	3580	230	6503.4727	2390	1300	1690	338
8/8/2018	25.809212	1040	258.48212	388	4810	1870	4990	3050	3550	230	6379.8325	2330	1290	1670	332
8/9/2018	25.810205	1035	259.06747	374	4800	1960	4870	3160	3490	228	6355.1045	2400	1275	1670	328
8/10/2018	25.729832	990	257.98608	382	4890	1940	4950	3070	3500	230	6330.3765	2410	1290	1675	334
8/13/2018	25.769522	935	260.86316	358	4850	1805	4650	2850	3370	226	6132.5522	2340	1260	1595	326
8/14/2018	25.769522	935	262.87714	350	4160	1815	4510	2730	3350	224	6058.3682	2330	1220	1570	320
8/15/2018	25.174154	930	262.08347	354	4030	1840	4620	2920	3430	224	6058.3682	2340	1215	1540	318
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
1/21/2019	24.659192	825	268.38558	368	4470	2470	4840	2250	4030	296	7500	2580	1605	1700	340
1/22/2019	24.138609	850	263.87708	358	4420	2460	4940	2270	4000	322	7575	2570	1615	1705	338
1/23/2019	24.163498	875	267.39783	366	4390	2450	5025	2320	3920	354	7650	2570	1600	1720	342
1/24/2019	24.579664	860	264.90439	358	4390	2390	4870	2330	3860	358	7775	2700	1600	1700	348
1/25/2019	24.915184	890	271.96582	358	4370	2420	4890	2190	3880	362	7850	3100	1615	1710	368
1/28/2019	24.773808	880	269.30283	360	4350	2400	4860	2130	3780	394	7700	3240	1610	1675	354
1/29/2019	24.787746	840	271.99573	354	4380	2370	4880	2060	3840	386	7700	3400	1610	1595	356
1/30/2019	24.964966	820	296.78067	358	4300	2530	4820	2110	3860	404	7700	3280	1610	1615	358
1/31/2019	24.659192	845	268.38558	364	4310	2570	4920	2170	3900	454	7750	2950	1600	1630	356

Lampiran Output R

Plot dendrogram penggerombolan jarak *Correlation*



Plot dendrogram penggerombolan jarak *Frechet*



Plot dendrogram penggerombolan jarak *Dynamic Time Warping*

