

# Pembentukan Portofolio Saham Melalui Proses Clustering Kurva Harga Saham Hasil Cubic-Spline

Faturachman Nugraha Sasmita<sup>1</sup>, Deni Saepudin<sup>2</sup>, Annisa Aditsania<sup>3</sup>

<sup>1,2,3</sup>Fakultas Informatika, Universitas Telkom, Bandung

<sup>1</sup>faturachmanugraha@students.telkomuniversity.ac.id, <sup>2</sup>denisaepudin@telkomuniversity.ac.id,

<sup>3</sup>aaditsania@telkomuniversity.ac.id

---

## Abstrak

Pelaku transaksi saham seringkali mengalami kesulitan dalam menentukan waktu yang tepat untuk membeli atau menjual saham. Hal ini disebabkan karena waktu pembelian saham dapat menentukan keuntungan investasi pada suatu saham. Maka dari itu, diperlukan portofolio saham yang dapat mendiversifikasi harga saham sehingga dapat membantu pembeli maupun penjual saham dalam bertransaksi di pasar modal. Penelitian ini membahas mengenai pembuatan portofolio saham melalui clustering kurva harga saham yang berasal dari metode *cubic spline*. *Cubic spline* untuk menginterpretasikan data yang sudah direduksi. Metode *clustering* pada penelitian kali ini dipakai untuk mengelompokkan koefisien *cubic-spline* dan menghasilkan 2,3, dan 4 clustering saham yang pengelompokkannya digunakan dengan metode K-means. Selanjutnya dilakukan pembentukan portofolio saham dengan memilih satu perwakilan dari setiap clustering berdasarkan rata-rata return setiap saham. Penelitian ini menghasilkan portofolio dengan nilai risiko terendah untuk pembagian cluster menjadi 4 cluster sebesar 0.0598 jika dibandingkan dengan pembagian cluster menjadi 2 cluster sebesar 0.1049 dan 3 cluster sebesar 0.2396.

**Kata kunci :** portofolio, saham, *cubic-spline*, *k-means*, *clustering*

---

## Abstract

Stock transaction participants often have difficulty in determining the right time to buy or sell shares. This is because the time to buy shares can determine the investment returns on a stock. Therefore, a stock portfolio is needed that can diversify stock prices so that it can help buyers and sellers of shares in transacting in the capital market. This study discusses the making of a stock portfolio through a stock price curve clustering derived from the cubic spline method. Cubic spline to interpret data that has been reduced. The clustering method in this study was used to classify the cubic-spline coefficient and produce 2,3, and 4 stock clustering whose classification is used by the K-means method. Next is the establishment of a stock portfolio by choosing one representative from each clustering based on the average return of each stock. This study produces a portfolio with the lowest risk value for the distribution of clusters into 4 clusters of 0.0598 when compared to the division of clusters into 2 clusters of 0.1049 and 3 clusters of 0.2396.

**Keywords:** portfolio, stock, *cubic-spline*, *k-means*, *clustering*

---

## 1. Pendahuluan

Saham adalah produk pada pasar modal yang naik turunnya tidak dapat dipastikan. Masalah yang dihadapi para investor adalah waktu yang tepat untuk membeli atau menjual saham agar mendapatkan keuntungan pada investasi saham. Pembentukan portofolio dilakukan agar harga saham terdiversifikasi, pembentukan portofolio menggunakan metode *clustering* dengan cara mengambil perwakilan dari masing masing klaster.

Metode *clustering* adalah alat yang sangat populer dalam analisis data di berbagai bidang penelitian seperti biologi, keuangan, dan domain aplikasi lainnya. Banyak algoritma *clustering* telah diperkenalkan dalam literatur. Satu kemungkinan klasifikasi metode ini dapat dilakukan sesuai dengan karya Han dan Kamber (2001). Menurut klasifikasi mereka metode utama adalah partisi, hierarkis, berbasis kepadatan, berbasis grid dan berbasis model.

Sebagian besar prosedur *clustering* tidak mempertimbangkan missing value dari suatu data. Selain itu, clustering memerlukan upaya komputasi ketika sejumlah besar pengukuran dipertimbangkan. Dengan tujuan mengatasi masalah masalah tersebut, teknik baru yang disebut "*parsimonious time series clustering*" baru baru ini diperkenalkan oleh Iorio, Frasso, D'Ambrosio, and Siciliano (2016)[6]. Teknik tersebut menggunakan metode *spline* yang bertujuan untuk membuat kurva lebih smooth, selain itu *spline* juga dapat mengatasi missing value atau data yang hilang.

Dalam membuat portofolio saham membutuhkan data yang banyak dan juga tidak sedikit kita menemukan data saham yang hilang dan itu akan menyulitkan saat membuat portofolio saham, maka dari itu dibutuhkan metode untuk mengatasi masalah tersebut. Penelitian ini membahas tentang pembentukan portofolio saham dengan mengelompokkan koefisien kubik-spline, dengan menggunakan koefisien hasil *cubic-spline* maka kurva yang

dimiliki lebih smooth dan tidak ada data yang hilang. Pendekatan berbasis *clustering spline* berkaitan juga dengan pengamatan data yang hilang yang memastikan kinerja pengelompokan berlangsung dengan baik. Untuk membuat portofolio bisa diambil perwakilan dari masing masing cluster untuk menjamin diversifikasi.

### Topik dan Batasannya

Pada penelitian ini akan dibahas cara mereduksi data pada harga saham dan membentuk portofolio saham melalui proses *clustering* kurva harga saham hasil *cubic-spline* dimana, data saham yang digunakan berjumlah 20 saham dalam jangka waktu 10 tahun.

### Tujuan

Tujuan dari penelitian ini untuk membentuk portofolio saham dengan nilai variansi portofolio terendah melalui proses *clustering* koefisien *cubic-spline* menggunakan metode *K-means*, data diambil dari harga saham selama 10 tahun.

## 2. Studi Terkait

### 2.1 Portofolio

Secara umum portofolio adalah kumpulan dokumen yang bertujuan untuk mendokumentasikan perkembangan suatu proses dalam mencapai tujuan yang telah ditetapkan. Portofolio saham adalah kumpulan aset investasi berupa saham, baik yang dimiliki perorangan atau perusahaan. *Equally Weighted* adalah model untuk membangun portofolio dengan cara membagi rata bobot dari saham. Berhubungan dengan masalah pemilihan portofolio, Benartzi dan Thaler (2001) menunjukkan bahwa seorang investor mengalokasikan kekayaannya berdasarkan gagasan naif tentang diversifikasi, yang dilakukan dengan menggunakan apa yang disebut aturan  $1/S$ . Strategi ini dikenal "investasi dengan bobot yang sama (EW)" [7]. Sesuai dengan strategi EW, investor hanya membagi kekayaan mereka di antara aset S yang ditawarkan, tidak memperhitungkan lebih canggih sebagai metodologi alokasi yang ditetapkan, seperti optimasi MV. Portofolio EW banyak digunakan oleh para praktisi pasar, karena konstruksinya tidak bergantung pada *expected return* dan juga terbukti sangat efisien.

### 2.2 K-Means

K-Means merupakan metode pengelompokan data dengan sistem partisi, metode ini mengelompokkan data yang ada ke dalam beberapa kelompok, dimana data dalam satu kelompok mempunyai karakteristik yang sama satu dengan yang lainnya dan mempunyai karakteristik yang berbeda dengan data yang ada di dalam kelompok yang lain. Input dari algoritma pengelompokan *K-Means* adalah jumlah kelompok "K" dan kumpulan data. Algoritma dimulai dengan perkiraan awal untuk "K" centroid yang dapat dihasilkan secara acak atau dipilih secara acak dari sekumpulan data. Centroid adalah titik pusat data, dalam hal ini kita mengasumsikan rata rata vektor sebagai centroid. Oyelade (2010) menyatakan algoritma ini disusun atas dasar ide yang sederhana. Sebaran obyek dan elemen pertama dalam cluster dapat dipilih untuk dijadikan sebagai titik tengah (centroid point) cluster[8]. Algoritma metode KMeans selanjutnya akan melakukan pengulangan langkah-langkah berikut sampai terjadi kestabilan (tidak ada obyek yang dapat dipindahkan):

- Langkah *assignment* data  
Setiap centroid mendefinisikan salah satu kelompok. Dalam langkah ini, setiap titik data ditetapkan ke centroid terdekat, menggunakan rumus *Euclidean Distance*

$$D(a, b) = \sqrt{\sum_{i=1}^n (b_i - a_i)^2} \quad (1)$$

Keterangan :

$D$  =jarak antara data  $a$  dan  $b$

Duran dan Odell (1974) menyatakan jika  $D(a, b)$  semakin kecil, kesamaan antara dua unit pengamatan semakin dekat. Syarat menggunakan jarak Euclid adalah jika semua fitur dalam dataset tidak saling berkorelasi. Jika terdapat fitur yang berkorelasi maka menggunakan konsep jarak Mahalanobis[10]. Agusta (2007) menyatakan kelanjutan dari jarak tersebut dicari yang terdekat sehingga data akan mengelompok berdasarkan centroid yang paling dekat[9].

- Langkah pembaruan centroid  
Pada langkah ini, centroid akan dihitung ulang, ini dilakukan dengan mengambil rata – rata semua titik yang ada pada centroid tersebut, iterasi akan berhenti apabila centroid yang baru sudah sama dengan centroid yang sebelumnya, atau tidak ada pergerakan data terhadap centroid, atau iterasi sudah mencapai maksimum iterasi.

Liao, Ho, dan Lin (2008) mengusulkan pendekatan dua langkah yang terdiri dari berlaku sebelum aturan asosiasi untuk mengusulkan kategori saham, algoritma k-means untuk menemukan portofolio saham yang dapat menyarankan saran investasi[11]. Teknik optimisasi portofolio dengan menggunakan dua matriks koefisien korelasi yang disaring berdasarkan prosedur pengelompokan hierarki diperkenalkan oleh Tola, Lillo, Gallegati, dan Mantegna (2008)[12]. Zhang dan Maringer (2009) mengusulkan metode evolusi berbasis populasi, yaitu Evolusi Diferensial, untuk mengidentifikasi partisi cluster optimal yang meningkatkan kinerja yang disesuaikan dengan risiko portofolio. Untuk memilih portofolio saham, Nanda, Mahanty, dan Tiwari (2010) melakukan tiga teknik pengelompokan yang terkenal (seperti k-means, SOM dan fuzzy c means) pada pengembalian keuangan dan pengembalian pasar[13]. Dengan cara yang berbeda, Silva dan Marques (2010) mengusulkan untuk melakukan dua analisis cluster. Yang pertama, milik metode hirarkis, untuk mendapatkan fitur cluster. Yang kedua, metode pengelompokan fitur berdasarkan SOM, untuk mendeteksi saham dengan perilaku serupa[14].

### 2.3 Cubic Spline

Pendekatan polinomial piecewise-paling umum menggunakan polinomial kubik antara setiap pasangan node berturut-turut dan disebut interpolasi spline kubik. Polinomial kubik umum melibatkan empat koefisien, sehingga ada fleksibilitas yang cukup dalam prosedur spline kubik untuk memastikan bahwa interpolant tidak hanya terdiferensiasi secara berkesinambungan pada interval, tetapi juga memiliki turunan kedua kontinu. Ide esensial untuk memasang data yang belum diketahui fungsinya dengan fungsi piecewise seperti rumus :

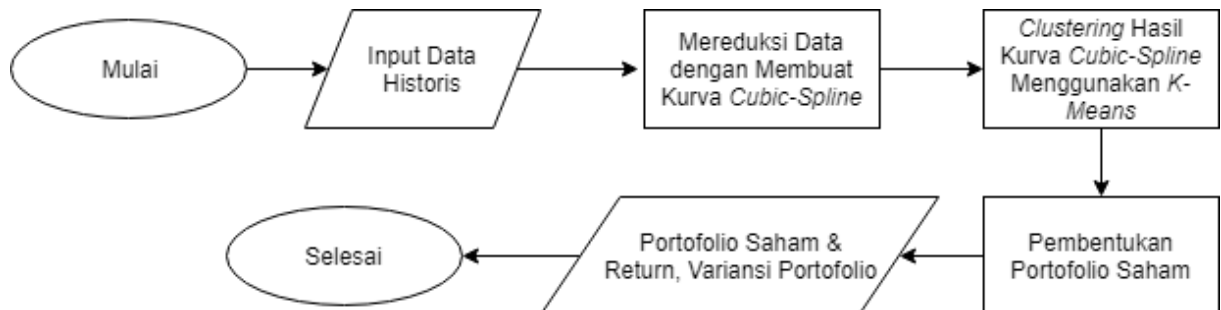
$$S(x) = \begin{cases} s_1(x) & \text{if } x_1 \leq x < x_2 \\ s_2(x) & \text{if } x_2 \leq x < x_3 \\ \vdots & \\ s_{n-1}(x) & \text{if } x_{n-1} \leq x < x_n \end{cases} \quad (2)$$

Dimana  $s_i$  adalah polinomial pangkat tiga yang di definisikan sebagai :

$$S_i(x) = a_i(x - x_i)^3 + b_i(x - x_i)^2 + c_i(x - x_i) + d_i \quad (3)$$

Untuk  $i = 1, 2, \dots, n - 1$

### 3. Sistem yang Dibangun



Gambar 3.1 Rancangan Sistem Implementasi Pembentukan Portofolio Saham

Penelitian ini untuk mempermudah investor dalam menilai saham menguntungkan atau tidak dengan menerapkan metode K-Means berdasarkan kurva hasil *cubic-spline* untuk pembentukan portofolio saham.

- **Data**

Data yang digunakan pada penelitian ini yaitu data harga saham mingguan dari 20 saham yang tergabung dalam LQ45 dengan rentang waktu 11 tahun (2005-2015) sebanyak 573 record. Berikut daftar saham yang dipilih :

Tabel 3.1 Saham yang dipakai pada penelitian.

Astra Agro Lestari (AALI.JK)	Adhi Karya (ADHI.JK)	Astra Internasional (ASII.JK)	Bank Central Asia (BBCA.JK)	Bank Negara Indonesia (BBNI.JK)
Bank Rakyat Indonesia (BBRI.JK)	Bank Mandiri (BMRI.JK)	Delta Dunia Makmur (DOID.JK)	Gudang Garam (GGRM.JK)	PT Hanjaya Mandala Sampoerna (HMSP.JK)
Indofood CBP Sukses Makmur (ICBP.JK)	Indomobil Sukses Internasional (IMAS.JK)	Indofood Sukses Makmur (INDF.JK)	Indocement Tunggul Prakasa (INTP.JK)	PP London Sumatra Indonesia Tbk (LSIP.JK)
Perusahaan Gas Negara (PGAS.JK)	Bukit Asam Tbk (PTBA.JK)	Surya Citra Media (SCMA.JK)	Timah Tbk (TINS.JK)	United Tractors Tbk (UNTR.JK)

- **Cubic-Spline**

Kurva *Cubic-Spline* dibuat berdasarkan 5,9, dan 25 persamaan yang diambil dari masing masing saham sehingga membentuk suatu kurva tertentu dari data yang sudah tereduksi. Kurva dibuat dari data prediksi yang didapat dari perhitungan spline.

- **Clustering dengan K-Means**

Koefisien kurva hasil spline yang memiliki tingkat kemiripan tinggi dimasukan pada satu cluster yang sama, dan masing-masing cluster memiliki tingkat kemiripan yang rendah. Tingkat kemiripan berdasarkan jarak dari masing masing data ke titik pusat data (centroid). Iterasi berhenti saat iterasi mencapai  $10^{10}$  dan centroid/pergerakan data tidak berubah lagi.

- **Pembentukan Portofolio Saham**

Perwakilan saham dari masing masing cluster lalu dibentuk menjadi portofolio saham. Perwakilan masing masing saham dipilih berdasarkan nilai rata rata return yang tertinggi. Berikut rumus untuk mencari nilai return saham :

$$Return = \frac{Saham\ kemarin - saham\ hari\ ini}{saham\ kemarin} \quad (4)$$

- **Pengujian kinerja**

Menghitung nilai return dan nilai risiko dengan membagi rata nilai bobot pada portofolio saham yang sudah dibentuk. Berikut adalah langkah langkah

1. Hitung kovarian dari tiap tiap saham yang terpilih sebelumnya dengan menggunakan rumus kovarian

$$\frac{\sum_{i=1}^n (x_i - x_{avg})(y_i - y_{avg})}{(n - 1)} \quad (5)$$

Keterangan :

$x_i$  = return saham  $x$  hari ke  $i$

$x_{avg}$  = rata rata return saham  $x$

$y_i$  = return saham  $y$  hari ke  $i$

$y_{avg}$  = rata rata return saham  $y$

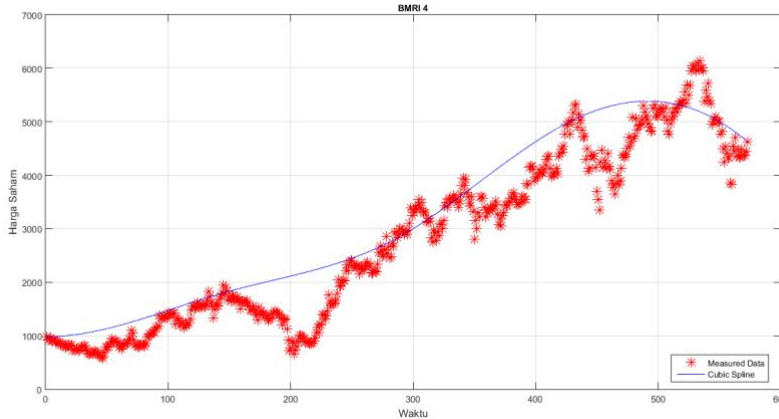
$n$  = jumlah data

2. Tetapkan bobot saham, karena penelitian ini menggunakan *equally weighted* maka bobot dibagi rata
3. Untuk menghitung nilai return lakukan perkalian matriks antara bobot saham dengan hasil kovarian dari saham.
4. Untuk menghitung nilai variansi lakukan perkalian matriks antara bobot saham (yang sudah di transpos) dengan hasil kovarian dari saham.

**4. Evaluasi**

**4.1 Hasil Pengujian**

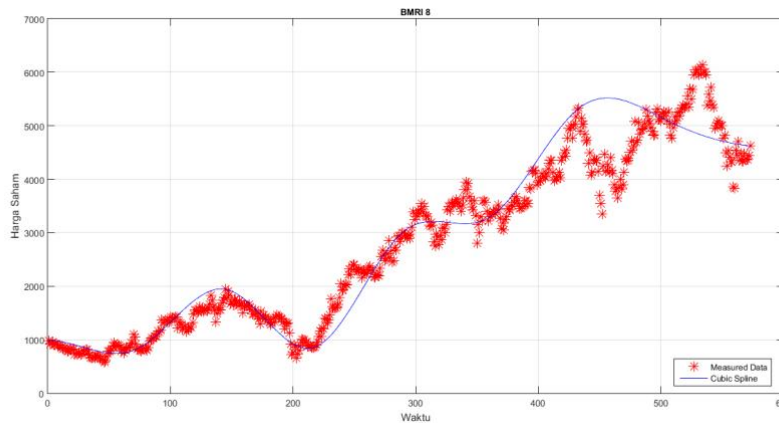
Pengujian spline kubik pada data harga saham BMRI diperoleh hasil sebagai berikut.



Gambar 4.1 Plot pola pergerakan 5 titik data saham BMRI spline kuadrat (biru) data actual (merah)

Proses pengujian spline kubik pada gambar 4.1 menggunakan 5 titik data saham (minggu ke-1, 144, 287, 430, 573) dan menghasilkan error sebesar  $8,404e + 08$ , menghasilkan 5 persamaan spline sebagai berikut

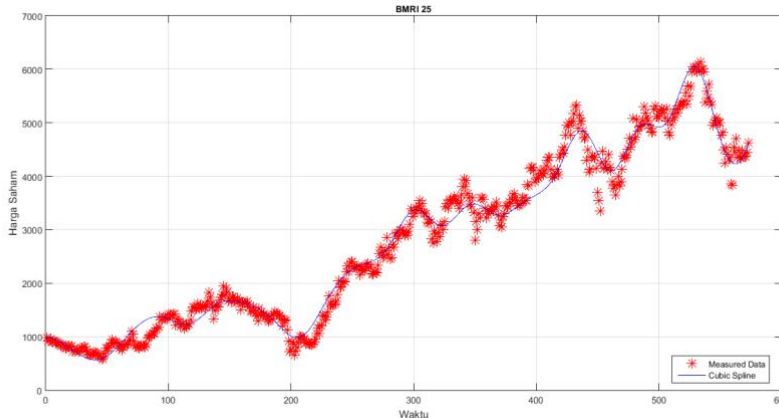
$$f(x) = \begin{cases} -0.0002x^3 + 0.0739x^2 - 0.1465x + 983.38 & 1 \leq x \leq 144 \\ 0.0002x^3 - 0.1249x^2 + 28.4851x - 390.94 & 145 \leq x \leq 287 \\ -0.0003x^3 + 0.3653x^2 - 112.2x + 13067.75 & 288 \leq x \leq 430 \\ -0.0001x^3 + 0.0981x^2 + 2.6693x - 3396.56 & 431 \leq x \leq 573 \end{cases} \quad (4)$$



Gambar 4.2 Plot pola pergerakan 9 titik data saham BMRI spline kuadrat (biru) data actual (merah)

Proses pengujian spline kubik pada gambar 4.2 menggunakan 9 titik data saham (minggu ke-1, 73, 145, 217, 289, 361, 433, 505, dan 573) dan menghasilkan error sebesar  $1,726e + 08$ , menghasilkan 9 persamaan spline sebagai berikut

$$f(x) = \begin{cases} 0.0029x^3 - 0.2503x^2 + 0.4830x + 983.06 & 1 \leq x \leq 73 \\ -0.0046x^3 + 1.4127x^2 - 120.92x + 3937.29 & 74 \leq x \leq 145 \\ 0.0063x^3 - 3.3758x^2 + 573.42x - 29622.66 & 146 \leq x \leq 217 \\ \vdots & \\ 0.000007x^3 - 0.0552x^2 - 14.25x + 16520.37 & 506 \leq x \leq 573 \end{cases} \quad (5)$$



Gambar 4.3 Plot pola pergerakan 24 titik data saham BMRI spline kuadratik (biru) data actual (merah)

Proses pengujian spline kubik pada gambar 4.2 menggunakan 24 titik data saham (minggu ke- 1, 25, 49, . . . , 553, dan 573) dan menghasilkan error sebesar  $1,477e + 08$ , menghasilkan 24 persamaan spline sebagai berikut

$$f(x) = \begin{cases} 0.01513x^3 - 0.8933x^2 + 1.6959x + 982.48 & 1 \leq x \leq 25 \\ 0.0102x^3 - 0.5377x^2 - 6.8391x + 1050.76 & 26 \leq x \leq 49 \\ -0.0177x^3 + 3.3981x^2 - 191.82x + 3948.903 & 50 \leq x \leq 73 \\ \vdots & \\ 0.0007x^3 - 1.1027x^2 + 433.18x - 48239.6 & 554 \leq x \leq 573 \end{cases} \quad (6)$$

Dari ketiga sample percobaan diatas, Saham dengan 24 titik akan digunakan untuk metode selanjutnya karena memiliki pola pergerakan spline yang mendekati dengan data aktualnya dan memiliki angka error terkecil jika dibandingkan dengan pola pergerakan spline 4 titik maupun 8 titik saham. Kemudian dari perhitungan spline akan didapatkan koefisien spline yang digunakan untuk menjadi input data dari pengelompokan data dengan menggunakan metode K-Means, pengelompokan data dibagi menjadi 2, 3, dan 4 kelompok.

Tabel 4.1 Pengelompokan saham dalam 2 cluster

Cluster 1	Cluster 2
AALI, GGRM, INTP	ADHI, ASII, BBKA, BBNI, BBRI, BMRI, ICBP, INDF, LSIP, UNTR, DOID, HMSP, IMAS, PGAS, PTBA, SCMA, TINS

Pengelompokan data saham diatas berdasarkan jarak data ke centroid, centroid tersebut diambil dari data secara random. Untuk membentuk portofolio setiap cluster akan dipilih satu saham untuk mewakili setiap cluster, pemilihan saham berdasarkan nilai rata – rata return dari setiap saham. Berikut perwakilan saham dari masing masing cluster :

- AALI (Perwakilan dari Cluster 1)
- PTBA (Perwakilan dari Cluster 2)

Tabel 4.2 Pengelompokan saham dalam 3 cluster

Cluster 1	Cluster 2	Cluster 3
ADHI, ASII, BBKA, BBNI, BBRI, BMRI, ICBP, INDF, LSIP, UNTR	AALI, GGRM, INTP	DOID, HMSP, IMAS, PGAS, PTBA, SCMA, TINS

Pengelompokan data saham diatas berdasarkan jarak data ke centroid, centroid tersebut diambil dari data secara random. Untuk membentuk portofolio setiap cluster akan dipilih satu saham untuk mewakili setiap cluster, pemilihan saham berdasarkan nilai rata – rata return dari setiap saham. Berikut perwakilan saham dari masing masing cluster :

- ICBP (Perwakilan dari Cluster 1)
- AALI (Perwakilan dari Cluster 2)
- PTBA (Perwakilan dari Cluster 3)

Tabel 4.3 Pengelompokan saham dalam 4 cluster

Cluster 1	Cluster 2	Cluster 3	Cluster 4
AALI, GGRM, INDF	ADHI, ASII, BBCA, BBNI, BMRI, IMAS, UNTR	BBRI, ICBP, INTP, LSIP, PGAS, SCMA	DOID, PTBA, TINS

Pengelompokan data saham diatas berdasarkan jarak data ke centroid, centroid tersebut diambil dari data secara random. Untuk membentuk portofolio setiap cluster akan dipilih satu saham untuk mewakili setiap cluster, pemilihan saham berdasarkan nilai rata – rata return dari setiap saham. Berikut perwakilan saham dari masing masing cluster :

- AALI (Perwakilan dari Cluster 1)
- IMAS (Perwakilan dari Cluster 2)
- SCMA (Perwakilan dari Cluster 3)
- PTBA (Perwakilan dari Cluster 4)

Lalu menghitung nilai return dan nilai risiko dari portofolio saham yang sudah dibentuk.

Tabel 4.4 Nilai bobot dari saham yang terpilih untuk 2,3 dan 4 cluster

2 Cluster		3 Cluster		4 Cluster	
Saham	Bobot	Saham	Bobot	Saham	Bobot
ICBP	0.5	ICBP	0.33	AALI	0.25
AALI	0.5	AALI	0.33	IMAS	0.25
		PTBA	0.33	SCMA	0.25
				PTBA	0.25

Pembagian 2 cluster ditetapkan 0.5 bobot (tabel kiri), untuk pembagian 3 cluster ditetapkan 0.33 bobot (tabel tengah), pembagian 4 cluster ditetapkan 0.25 bobot (tabel kanan) untuk masing masing saham.

Tabel 4.6 Hasil Perhitungan Return Portofolio dan Variansi Portofolio

2 Cluster		3 Cluster		4 Cluster	
Return Portofolio	0.0984	Return Portofolio	0.0671	Return Portofolio	0.0532
Variansi Portofolio	0.2396	Variansi Portofolio	0.1049	Variansi Portofolio	0.0598

## 5. Kesimpulan

Spline kubik mampu merepresentasikan pola pergerakan saham yang sudah direduksi dengan lebih *smooth*. K-Means clustering berperan untuk mengelompokkan saham dengan menggunakan koefisien dari spline menjadi tiga buah kelompok lalu dipilih data dengan nilai return tertinggi sebagai perwakilan dari tiga kelompok untuk menjamin diversifikasi. Dengan membuat penyusunan portofolio saham dengan menggunakan metode *equally weighted* investor dapat mengetahui nilai bobot dari saham tersebut serta dapat mengetahui return dan variansi dari portofolio tersebut. Dari hasil perhitungan pada penelitian ini berdasarkan variansi portofolio terendah dapat dipilih dari pembagian empat cluster dengan variansi portofolio sebesar 0.0598.

## Daftar Pustaka

- [1] P. Eilers and B. Marx, "Flexible with B-splines and penalties," *Statistical Science*, pp. 89-102, 1996.
- [2] J. Han, M. Kamber and J. Pei, *Data Mining Concept and Techniques Third Edition*, Elsevier, 2011.
- [3] H. Markowitz, "Portfolio Selection," *The Journal of Finance*, pp. 77-91, 1952.
- [4] R. Burden and D. Faires, *Numerical Analysis 9th Edition*, Brooks: Cengage Learning, 2011.
- [5] "Informatikalogi," 2019. [Online]. Available: <https://informatikalogi.com/algorithm-k-means-clustering/>. [Accessed 5 Juli 2019].
- [6] C. Iorio, G. Frasso, A. D'Ambrosio and R. Siciliano, "Parsimonious time series clustering using p-splines," *Expert Systems with Applications*, vol. 52, pp. 26-38, 2016.
- [7] S. Benartzi and R. H. Thaler, "Naive Diversification Strategies in Defined Contribution Saving Plans," *American Economic Review*, vol. 91, no. 1, pp. 79-98, 2001.
- [8] O. J. O. O. O. O. and I. C. O. , "Application of k Means Clustering algorithm for prediction of Students Academic Performance," *International Journal of Computer Science and Information Security*, vol. 7, pp. 292 - 295, 2010.
- [9] Y. Agusta, "K-Means-Penerapan, Permasalahan, dan Metode Terkait," *Jurnal Sistem dan Informatika*, pp. 47-60, 2007.
- [10] B. S. Duran and P. L. Odell, *Cluster Analysis*, Texas: Springer-Verlag, 1974.
- [11] S. H. Liao, H. H. Ho and H. W. Lin, "Mining stock category association and cluster on Taiwan stock market," *Expert Systems with Applications*, vol. 1, no. 35, pp. 19-29, 2008.
- [12] V. Tola, F. Lillo, M. Gallegati and R. Mantegna, "Cluster analysis for portfolio optimization," *Journal of Economic Dynamics and Control*, vol. 32, no. 1, pp. 235-258, 2008.
- [13] S. R. Nanda, B. Mahanty and M. K. Tiwari, "Clustering Indian stock market data for portfolio management," *Expert Systems with Applications*, vol. 37, no. 12, pp. 8793-8798, 2010.
- [14] B. Silva and N. C. Marques, "Feature Clustering with Self-Organizing Maps and an Application to Financial Time-Series Portfolio Selection," in *International Conference on Neural Computation (ICNC) -* , Valencia, 2010.