

Penerapan K-Means Clustering dengan Menggunakan Metode Elbow dan Silhouette Coefficient pada Bank Berdasarkan Suku Bunga

Lembaga Penjamin Simpanan (LPS) dalam melakukan penjaminan simpanan masyarakat menerapkan tiga syarat simpanan yang layak bayar oleh LPS yaitu tercatat pada pembukuan bank, tingkat bunga yang diterima tidak melebihi tingkat bunga penjaminan LPS, tidak melakukan tindakan yang merugikan bank. Saat ini terdapat beberapa bank yang sudah mematuhi tingkat penjaminan LPS dan beberapa belum menerapkan hal tersebut. Oleh karena itu, perlu pengelompokkan bank berdasarkan suku bunga baik suku bunga tertinggi maupun suku bunga terendah dari suatu bank untuk melihat seberapa patuh bank tersebut terhadap tingkat bunga penjaminan LPS agar mempermudah analisis dalam mapping bank yang tidak patuh terhadap LPS rate. Jangka panjang dari analisis ini yaitu untuk pengawasan bank dan himbauan yang lebih intensif pada bank yang memiliki bunga melebihi LPS rate. Pengelompokkan dapat dilakukan dengan dataset Laporan Harian Bank Umum (LHBU) dan menggunakan algoritma K-Means Clustering dimana proses pengelompokkan ini berdasar pada karakteristik atau ciri-ciri data yang mirip. Pada proses pengelompokkan ini dilakukan pula perbandingan pemilihan nilai k pada K-Means Clustering menggunakan Elbow Method dan Silhouette Coefficient dapat dilakukan untuk mempermudah pemilihan berdasarkan dataset yang sudah ada.

Pendahuluan

Dalam melakukan penjaminan simpanan masyarakat, LPS menerapkan tiga syarat simpanan yang layak bayar oleh LPS yaitu:

1. Tercatat pada pembukuan bank;
2. Tingkat bunga yang diterima tidak melebihi tingkat bunga penjaminan LPS;
3. Tidak melakukan tindakan yang merugikan bank.

Saat ini terdapat 107 (seratus tujuh) Bank Umum di Indonesia. Sudah banyak bank yang menaati tingkat penjaminan yang ditetapkan oleh LPS. Berdasarkan data pada bulan Mei 2022, tingkat bunga deposito Bank BRI adalah sebesar 3%, Bank Mandiri sebesar 2.6%, Bank BCA sebesar 1.9%, dan Bank BNI sebesar 2.5%. Sedangkan, terdapat beberapa bank digital yang memiliki tingkat bunga penjaminan lebih dari yang ditentukan LPS sebesar 3.5% (dapat dilihat di website LPS - <https://lps.go.id/>). Dilihat dari hal tersebut, terdapat beberapa bank yang sudah mematuhi tingkat penjaminan LPS dan beberapa belum menerapkan hal tersebut. Oleh karena itu, perlu adanya sebuah sistem pengelompokan bank berdasarkan suku bunga baik suku bunga tertinggi maupun suku bunga terendah dari suatu bank untuk melihat seberapa patuh bank tersebut terhadap tingkat bunga penjaminan LPS. Pengelompokan ini dilakukan dengan menggunakan algoritma K-Means Clustering dimana nilai cluster atau nilai k dari K-Means Clustering ditentukan dengan menggunakan 2 metode yaitu Elbow Method dan Silhouette Coefficient. Penentuan nilai k ini bertujuan untuk membandingkan keluaran K-Means Clustering dengan nilai k mana yang paling baik dan dapat di analisis.

Tujuan dari analisis ini adalah untuk mempermudah mapping atau pengelompokan bank yang tidak patuh terhadap LPS rate. Jangka panjang dari analisis ini yaitu untuk pengawasan bank dan himbauan yang lebih intensif pada bank yang memiliki bunga melebihi LPS rate dengan menggunakan algoritma K-Means Clustering. Tahapan dalam melakukan analisis ini yakni, menganalisa kebutuhan data, pengumpulan data, cleansing data dan analisis menggunakan algoritma K-Means.

Referensi

Tingkat Bunga Penjaminan LPS

Tingkat bunga penjaminan LPS merupakan tingkat bunga yang menjadi acuan untuk simpanan nasabah. Hal ini dapat diartikan sebagai batas atas tingkat bunga yang dapat diberikan bank kepada nasabah untuk dapat memenuhi kriteria penjaminan. Sehingga, suku bunga simpanan di bank tidak boleh melebihi tingkat bunga penjaminan LPS agar simpanan bisa dijamin oleh LPS ketika bank dilikuidasi. Penetapan tingkat suku bunga ini mempertimbangkan dinamika pergerakan suku bunga pasar simpanan Rupiah dan Valas, serta analisa terhadap perkembangan berbagai faktor seperti pemulihan ekonomi, kondisi pasar keuangan, kinerja perbankan khususnya fungsi intermediasi dan prospek likuiditas. Suku bunga penjaminan ini diumumkan secara berkala oleh LPS untuk terus dapat menyesuaikan dengan kondisi perekonomian dan kebijakan pemerintah secara umum. Penetapan tingkat bunga penjaminan terbaru diputuskan untuk tetap dipertahankan yaitu 3,5% untuk Bank Umum dan 0,25% untuk valuta asing di Bank Umum, serta 6,00% untuk di BPR.

Tingkat Bunga Penjaminan	
Periode: 28-05-2022 - 30-09-2022	
Bank Umum	
IDR 3.50%	Valas 0.25%
BPR	
IDR 6.00%	

Gambar 1 Tingkat Bunga Penjaminan LPS
Sumber: (Lembaga Penjamin Simpanan (LPS), Mei 2022)

Algoritma K-Means

K-Means merupakan salah satu metode dalam data mining yang dapat mengelompokkan data atau clustering sebuah data dalam bentuk satu cluster atau lebih cluster sehingga data dengan karakteristik yang sama dapat dikelompokkan dalam satu cluster. K-Means clustering adalah metode yang termasuk pada clustering non-hirarki. Pada dasarnya clustering adalah metode yang berfungsi untuk mengkategorikan sesuai dengan atribut yang sama atau karakteristik dengan data lainnya. Clustering merupakan suatu metode pada data mining yang dimana proses kerja algoritma ini sifatnya tanpa arahan (unsupervised). Hal ini berarti algoritma tersebut tidak memerlukan suatu training

1. Elbow Method

Elbow method adalah metode yang sering digunakan untuk menentukan jumlah cluster yang akan digunakan pada K-Means Clustering. Clustering sendiri bertujuan untuk meminimumkan jarak antara data poin dan centroid serta memaksimumkan jarak antara centroid yang ada menggunakan within-cluster sum of squares atau WCSS. Tujuan dari perhitungan ini adalah agar mempermudah interpretasi data. Metode ini merupakan salah satu metode yang bertujuan untuk menentukan jumlah cluster yang tepat melalui persentase hasil perbandingan antara jumlah cluster yang akan membentuk siku pada suatu titik. Jika nilai cluster pertama dengan nilai cluster kedua memberikan sebuah sudut dalam grafik atau nilainya mengalami penurunan paling besar, maka jumlah nilai cluster tersebut yang tepat

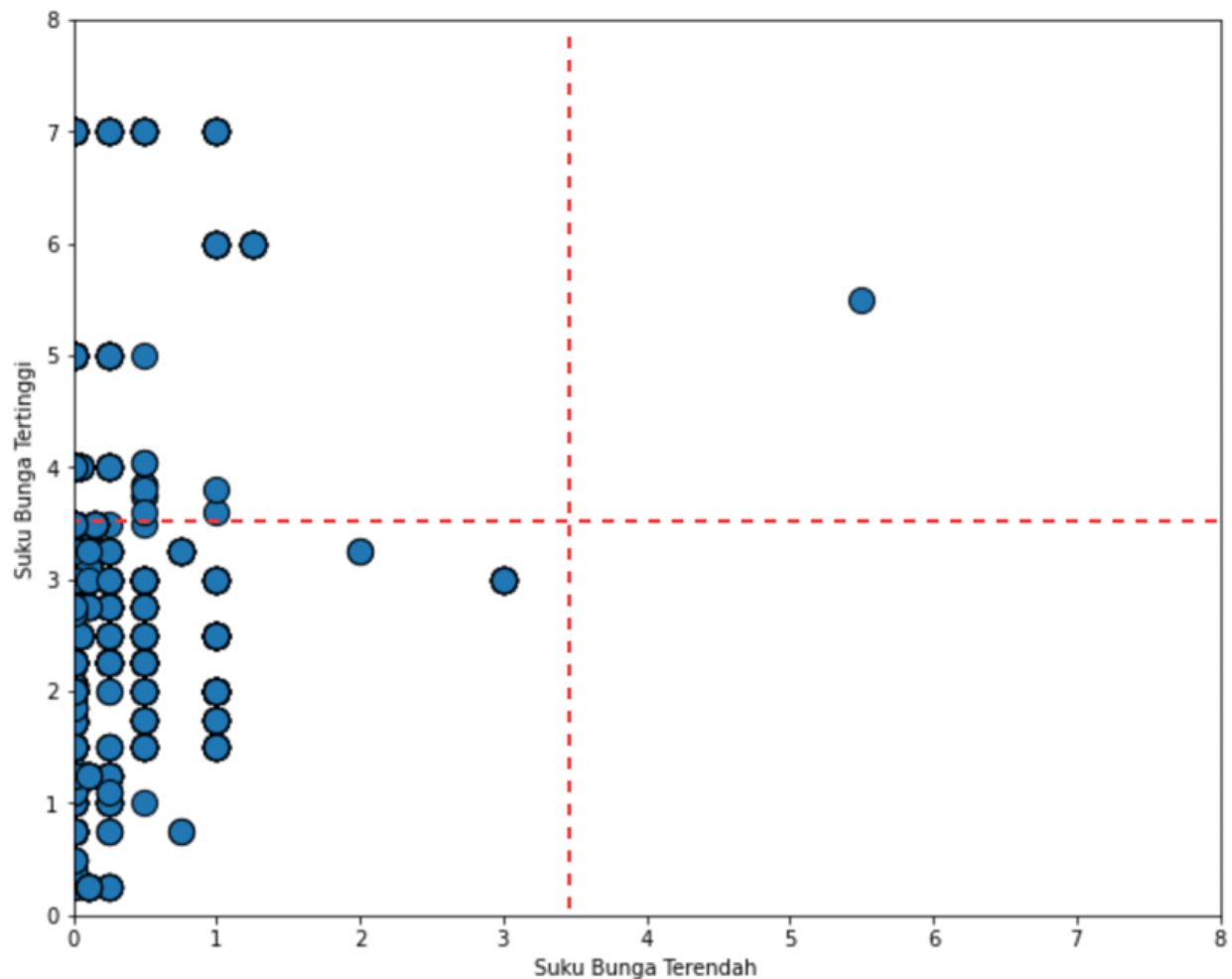
2. Silhouette Coefficient

Silhouette coefficient merupakan metode yang digunakan untuk melihat kualitas dan kekuatan cluster. Hal yang dilihat yaitu seberapa baik atau buruk suatu objek ditempatkan dalam suatu cluster. Metode ini merupakan metode gabungan dari separasi dan kohesi

Analisis Masalah

Dataset

Dataset yang digunakan merupakan data LHBK dimana data tersebut berisi laporan suku bunga tertinggi dan terendah dari bank umum selama 1 bulan dari 1 April 2022 sampai dengan 30 April 2022. Jumlah keseluruhan bank yang ada pada dataset yang digunakan sebanyak 85 bank dengan jumlah baris data sebanyak 1615 baris. Dataset ini terdiri dari beberapa kolom diantaranya yaitu tanggal pelaporan suku bunga bank, nilai suku bunga dari yang tertinggi sampai yang terendah dan rata-rata suku bunga masing-masing bank.



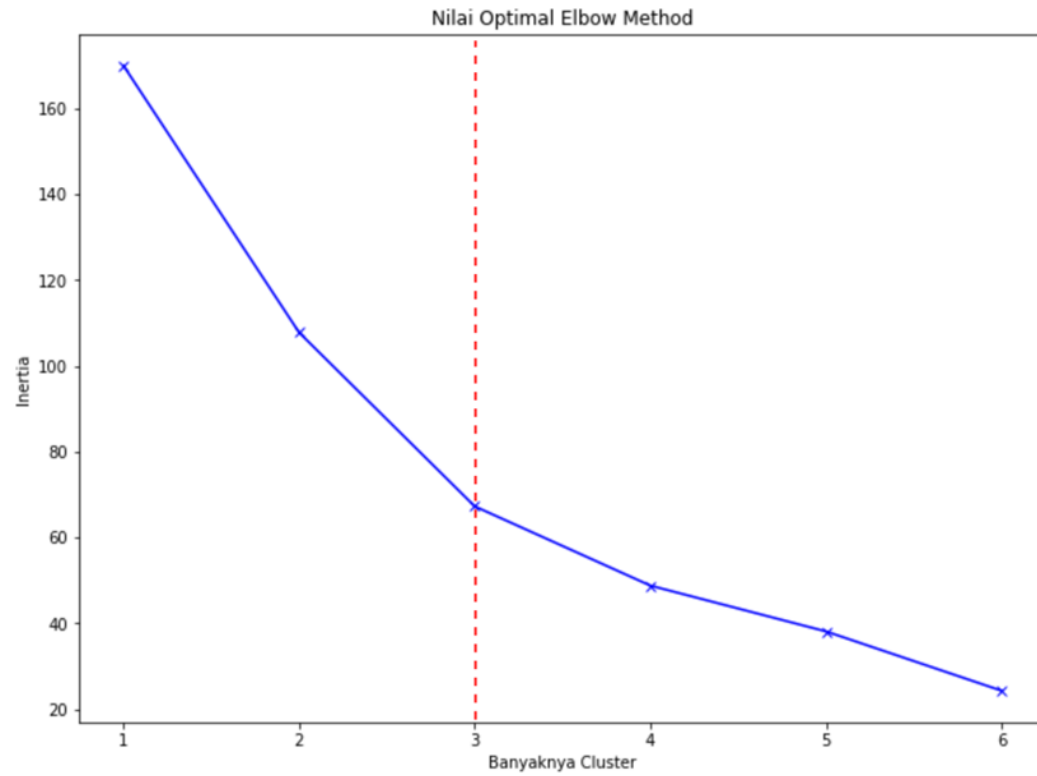
Gambar 2 Sebaran Bank Berdasarkan Suku Bunga

Dataset yang ada kemudian dilihat sebaran datanya melalui scatter plot berdasarkan suku bunga tertinggi dan suku bunga terendah yang tertera pada Gambar 2. Dari data terlihat bahwa terdapat beberapa bank yang memiliki nilai suku bunga melebihi tingkat bunga penjaminan LPS. Terdapat bank yang memiliki suku bunga tertinggi hingga 7%, bahkan terdapat pula bank yang memiliki suku bunga terendah mencapai hampir 6%.

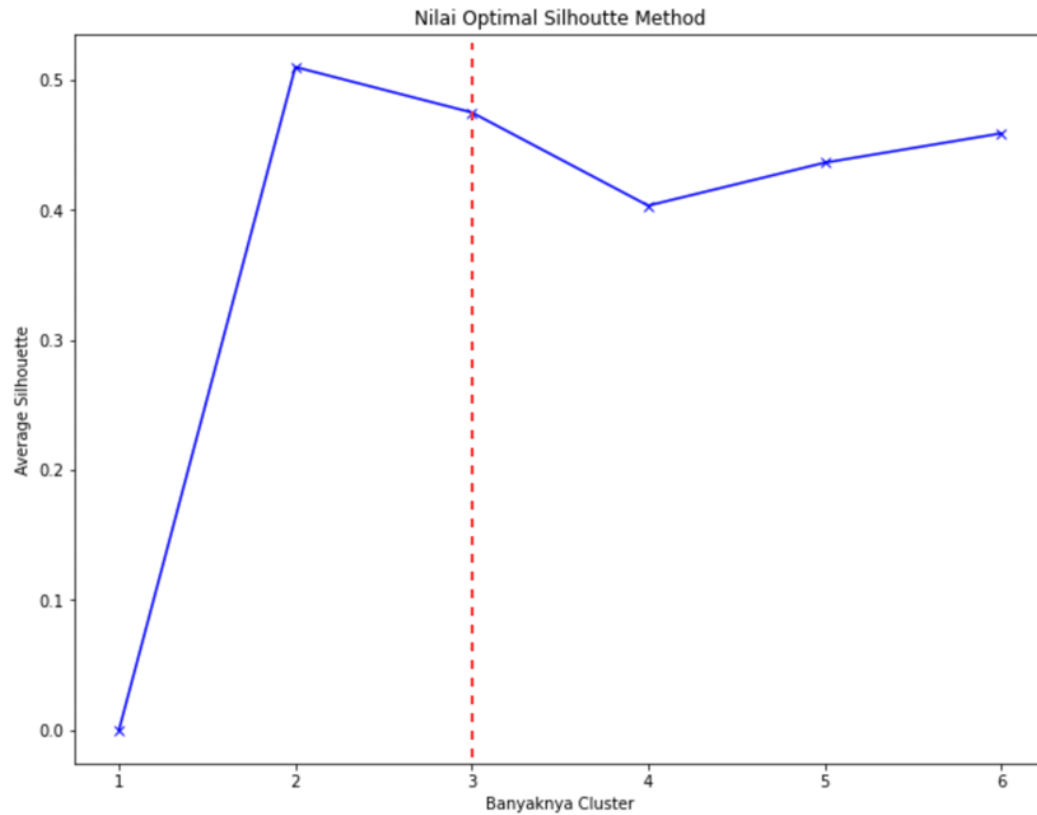
Penentuan Nilai Cluster (Nilai 'k')

Setelah melakukan pra proses data, tahapan selanjutnya adalah proses penentuan nilai cluster atau nilai k. Nilai cluster atau nilai k merupakan jumlah cluster atau kelompok yang optimal. Pada tahapan ini penentuan nilai k menggunakan 2 metode yaitu elbow method dan silhouette coefficient dimana nanti hasil clustering akan disesuaikan dengan nilai k yang didapat dari penentuan kedua metode. Penentuan nilai k menggunakan 2 metode ini bertujuan untuk membandingkan pola pengelompokkan pada saat clustering dengan algoritma K-Means.

Berdasarkan grafik nilai optimal elbow method yang tertera pada Gambar 3 di dapat nilai optimal k adalah 3. Kemudian sebagai perbandingan, dilakukan penentuan dengan menggunakan silhouette coefficient melalui grafik dan scoring. Grafik nilai optimal silhouette method yang tertera pada Gambar 4 juga menunjukkan bahwa nilai optimal k adalah 3.



Gambar 3 Grafik Nilai Optimal Elbow Method



Gambar 4 Grafik Nilai Optimal Silhouette Coefficient




Pada Tabel 1 merupakan hasil scoring silhouette coefficient dimana pada percobaan nilai k dimasukkan nilai 2 sampai dengan 6 untuk melihat hasil score terbaik pada percobaan dataset yang ada. Dari percobaan tersebut, nilai k yang optimal dan memiliki score silhouette terbesar adalah sebesar 0,545 dengan nilai k yaitu 3. Dilihat dari hasil elbow method dan juga silhouette coefficient, maka dapat diambil kesimpulan bahwa nilai k yang optimal untuk mengelompokkan cluster menggunakan algoritma K-Means adalah nilai k sebesar 3.

Tabel 1 Tabel Score Silhouette

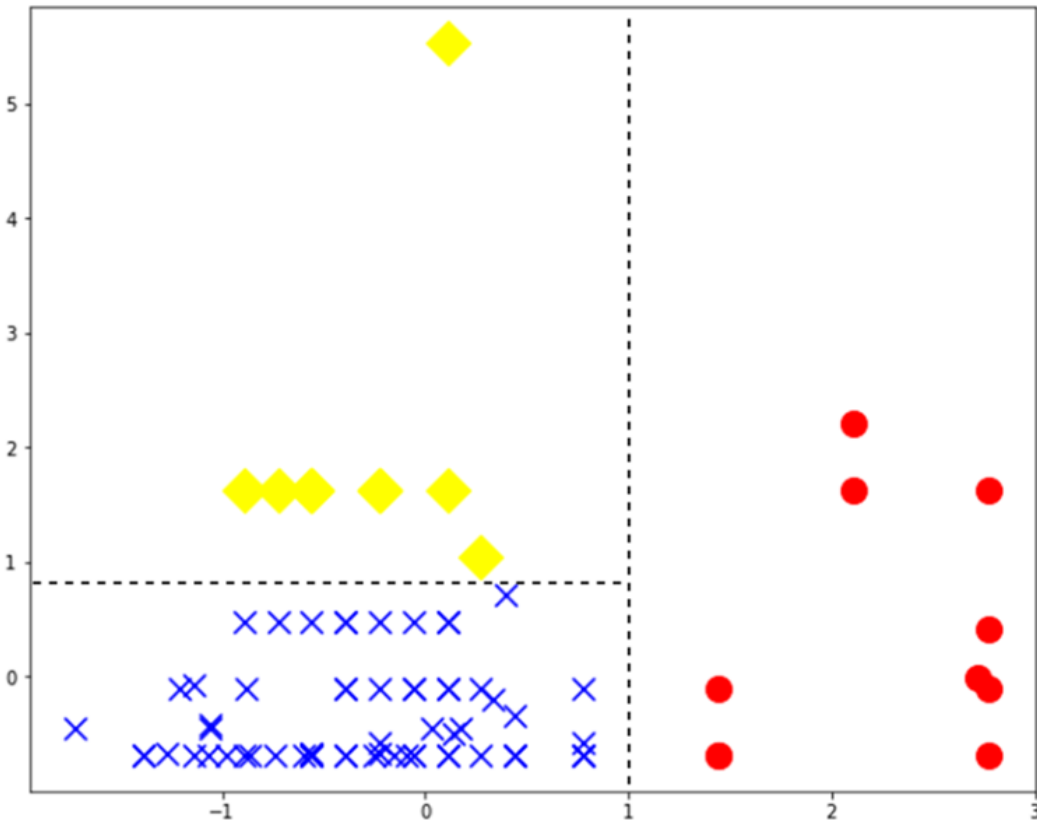
Nilai 'k'	Score Silhouette
2	0,510
3	0,545
4	0,403
5	0,444
6	0,460

Hasil clustering dengan menggunakan algoritma K-Means diperoleh hasil cluster sebanyak 3 total cluster dimana cluster ini terbentuk berdasarkan kemiripan dari masing-masing data yang ada. Hasil clustering ini menerapkan nilai 'k' sebesar 3 karena dari penentuan nilai cluster atau nilai 'k' yang sebelumnya menggunakan 2 metode, yakni elbow method dan silhouette coefficient. Keduanya menghasilkan nilai optimal yang sama yaitu 3 dengan nilai score silhouette coefficient sebesar 0,545.

Tabel 2 Pembagian Cluster

Cluster	Pola Cluster
1	
2	
3	

Dari pengelompokkan cluster tersebut dapat dilihat bahwa terdapat cluster yang memiliki suku bunga tertinggi lebih dari tingkat bunga penjaminan LPS. Data ini tergambar dengan pola bujur sangkar berwarna kuning dan masuk dalam cluster 3. Selain itu terdapat pula data yang masuk dalam cluster 1 dan cluster 2 dimana hasil reshape data suku bunga tertinggi bank tersebut tidak melebihi tingkat bunga penjaminan LPS. Perbedaan dari cluster 1 dan cluster 2 adalah terletak pada hasil reshape data suku bunga terendah bank dimana pada cluster 1 terdapat hasil data minus. Hal ini dapat terjadi ketika tahapan pra proses data menghasilkan data yang dibawah rata-rata suku bunga terendah.



Gambar 5 Hasil Clustering K-Means dengan Jumlah Cluster 3

Hasil menunjukkan bahwa kedua metode yang digunakan untuk penentuan nilai 'k' menghasilkan hasil yang sama. Kedua metode tersebut dapat digunakan secara bersamaan dengan tujuan untuk menguji dan memastikan apakah hasil dari penentuan nilai 'k' sudah benar dan optimal. Penentuan nilai 'k' menggunakan silhouette coefficient dapat melengkapi elbow method dikarenakan silhouette coefficient tidak hanya menampilkan nilai 'k' melalui grafik, melainkan terdapat pula pengecekan score nilai 'k' mana yang optimal untuk digunakan pada tahap clustering menggunakan algoritma K-Means.

Akan tetapi, penggunaan algoritma K-Means pada tahap clustering ini memiliki kekurangan dimana terdapat data hasil cluster yang memiliki jarak cukup jauh dengan data lainnya. Terdapat satu data poin pada cluster 3 yang jaraknya jauh dengan data poin lainnya pada cluster 3. Hal ini dapat diperbaiki dengan melakukan pengecekan data kembali dan melakukan pra proses data untuk memastikan bahwa jarak antar data poin dapat menggambarkan karakteristik dan ciri-ciri dari masing-masing data yang ada.

Dari hasil pengelompokkan tersebut, kemudian dilakukan pengecekan kembali dari dataset yang ada. Pengecekan ini bertujuan untuk melihat bank mana saja yang masuk dalam cluster 3 atau dapat dikatakan bank yang memiliki suku bunga tertinggi dan tidak patuh terhadap tingkat bunga penjaminan LPS.

Daftar Pustaka

- Dewi, D. A., & Pramita, D. A. (2019). Analisis Perbandingan Metode elbow Dan silhouette pada Algoritma clustering K-medoids dalam Pengelompokan Produksi Kerajinan Bali. *Matrix : Jurnal Manajemen Teknologi dan Informatika*, 9(3), 102-109.
- Febrianti, A. F. (2018). K-Means Clustering Dengan Metode Elbow Untuk Pengelompokkan Kabupaten dan Kota di Jawa Timur Berdasarkan Indikator Kemiskinan. *Jurnal Statistika Industrial Ekonomi*, 863-870.
- I. H. Witten, E. Frank, and M. A. Hall, *Data Mining: Practical Machine Learning Tools and Techniques*. Elsevier Science, 2011.
- M. A. Syakur, B. K. Khotimah, E. M. S. Rochman, and B. D. Satoto, "Integration K-Means Clustering Method and Elbow Method for Identification of the Best Customer Profile Cluster," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 336, no. 1, 2018.
- Nurdiawan, O., & Salim, N. (2018). Penerapan Data Mining Pada Penjualan Barang Menggunakan Metode Metode Naive Bayes Classifier Untuk Optimasi Strategi Pemasaran. *Jurnal Teknologi Informasi Dan Komunikasi*, 11(1), 84-95.
- Retnoningsih, E., & Pramudita, R. (2020). Mengenal Machine Learning Dengan Teknik Supervised dan Unsupervised Learning Menggunakan Python. *Bina Insani ICT Journal*, 7(2), 156-165.
- Wibowo, A., & Sasongko, R. (2022). Penerapan Data Mining Pada Suku Bunga Deposito di Indonesia Menggunakan Metode K-Means Clustering Untuk Pengelompokkan Profitabilitas ((Application of Data Mining on Deposit Investment Rate in Indonesia Using K-Means Clustering Method for Profitability Clustering). *ASCARYA: Islamic Science, Culture, and Social Studies*, 2(1), 70-80.