

IMPLEMENTASI *FUZZY C-MEANS* UNTUK *CLUSTERING* DATA HARGA SAHAM HARIAN PADA PT. ASTRA INTERNATIONAL TBK

Ridha Nafila Turrahma¹, Adde Nanda Caesario Putra², Muhammad Daffa Alfajri³,
Rizki Gusmanto⁴, Widhia KZ Oktoberza⁵

Program Studi Informatika, Fakultas Teknik, Universitas Bengkulu
Jl. WR. Supratman Kandang Limun Bengkulu 38371 Indonesia

nafilearidhaa@gmail.com, nandaadde@gmail.com, dfajri856@gmail.com,
rizkigusmanto16@gmail.com, widhiakz@unib.ac.id

Abstrak:

Pada metode *cluster* non-hirarki terdapat beberapa algoritma *clustering* data, salah satunya adalah metode *Fuzzy C-Means* (FCM). Pada proses pengelompokan (*clustering*), dalam metode *Fuzzy C-Means* (FCM) satu data bisa menjadi dua *cluster* atau lebih. Tujuan penelitian ini adalah untuk meng*cluster*kan data saham harian PT. Astra, Tbk dengan metode *Fuzzy C-Means* (FCM) secara teoritis serta mengetahui aplikasinya. Langkah awal dalam penelitian ini adalah mengetahui metode FCM secara teoritis kemudian mengaplikasikannya pada data saham harian PT. Astra, Tbk. Dan pada akhir penelitian ini diperoleh kesimpulan bahwa *Fuzzy C-Means* (FCM) merupakan metode *clustering* data yang digolongkan sebagai metode pengklasifikasian yang bersifat unsupervised (tanpa arahan). Pengalokasian ulang data ke dalam masing-masing *cluster*, Untuk penerapannya, peneliti berupaya mengelompokkan harga saham harian berdasarkan harga sebelum (open) dan harga tertinggi (high) dengan metode FCM. Hasil penelitian diperoleh menghasilkan 3 *cluster* yaitu '*cluster* rendah' dengan anggota pada bulan Januari - Maret rata-rata harian harga saham *open* Rp. 3835 dan *high* Rp. 3933, '*cluster* sedang', anggota pada bulan Maret - Mei dengan rata-rata harian harga saham *open* Rp. 4988 dan *high* Rp. 5061, dan '*cluster* tinggi' pada bulan Mei - Desember dengan anggota rata-rata harian harga saham *open* Rp. 6561 dan *high* Rp. 6648. Namun, terdapat 7 data yang tidak berkelompok sesuai dengan *range* bulan yang dianalisis. Penelitian ini diharapkan mampu menjadi bahan pertimbangan para investor untuk membeli saham di pasar modal.

Kata Kunci: *Clustering* Data, Metode *Fuzzy C-Means*, Harga Saham

Abstract: *In the non-hierarchical cluster method, there are several data clustering algorithms, one of which is the Fuzzy C-Means (FCM) method. In the clustering process, in the Fuzzy C-Means (FCM) method one data can become two or more clusters. The purpose of this research is to cluster the daily stock data of PT. Astra, Tbk with the theoretical Fuzzy C-Means (FCM) method and know its application. The first step in this research is to know the FCM method theoretically and then apply it to the daily stock data of PT. Astra, Tbk. And at the end of this study it was concluded that Fuzzy C-Means (FCM) is a data clustering method which is classified as an unsupervised classification method (without direction). Re-allocation of data into each cluster. For its application, the researcher seeks to classify daily stock prices based on the open price and the high price using the FCM method. The results of the research were*

obtained to produce 3 clusters namely 'low cluster' with members in January - March the average daily open price is Rp. 3835 and high Rp. 3933, 'medium cluster' members in March - May with an average daily open share price of Rp. 4988 and high Rp. 5061, and 'high cluster' in May - December with an average daily open price of Rp. 6561 and high Rp. 6648. However, there are 7 data that are not grouped according to the range of months being analyzed. This research can be considered for investor to put their money on market capital

Keyword: *Clustering* Data, *Fuzzy C-Means*, *Stock Prices*

I. PENDAHULUAN

Pasar modal dapat mempermudah masyarakat dan para investor untuk mengetahui perkembangan tentang harga saham setiap

perusahaan-perusahaan di Indonesia yang sudah go public. Salah satu perusahaan yang sudah terdaftar di bursa efek Indonesia dan telah go public adalah PT Astra Internasional Tbk. Masyarakat yang ingin mengetahui harga saham PT Astra Internasional Tbk setiap harinya dapat melihatnya di Bursa Efek Indonesia (BEI). Keterbukaan data harga saham yang dimiliki oleh PT Astra Internasional Tbk, hal ini dapat digunakan oleh masyarakat dan para investor untuk mendapatkan keuntungan. Namun, untuk mendapatkan keuntungan, maka masyarakat dan para investor harus mengetahui waktu yang tepat untuk melakukan penjualan dan pembelian saham yang diinginkan di PT Astra Internasional Tbk.

Harga Saham adalah harga penutup yang sering ditampilkan di bursa efek Indonesia atau di pasar modal Indonesia yang menampilkan grafik atau harga yang dapat digunakan masyarakat, investor dan perusahaan-perusahaan yang ingin melakukan penjualan dan pembelian saham, sehingga memberikan keuntungan secara pribadi. Ada tiga jenis penilaian ekuitas: nilai nominal, nilai pasar intrinsik, dan nilai efektif. Nilai nominal adalah harga yang dikutip dari saham terdapat, nilai intrinsik adalah nilai saham pada saat diperdagangkan, dan nilai efektif adalah nilai saham dengan kurs resmi pada saat saham diperdagangkan di bursa. bursa efek Estimasi harga (Sudirman, 2015).

Pembentukan *cluster* atau kelompok data merupakan salah satu teknik yang digunakan dalam mengekstrak pola kecenderungan suatu data. Analisis *cluster* atau *clustering* merupakan proses membagi data dalam suatu himpunan ke dalam beberapa kelompok yang kesamaan datanya dalam suatu kelompok lebih besar daripada kesamaan data tersebut dengan data dalam kelompok lain.

Metode *Fuzzy C-Means* yang digunakan dipilih karena metode ini adalah suatu teknik pengelompokan data dimana keberadaan tiap-tiap titik data dalam suatu *cluster* ditentukan oleh derajat keanggotaan. *Fuzzy C-Means* adalah algoritma pengelompokan yang terawasi, karena pada algoritma *Fuzzy C-Means* jumlah *cluster* yang akan dibentuk perlu diketahui terlebih dahulu. Konsep dasar algoritma *Fuzzy C-Means* adalah menentukan pusat kelompok yang akan menandai lokasi rata-rata untuk tiap-tiap *cluster*. Pada kondisi awal, pusat *cluster* ini masih belum akurat. Tiap-tiap titik data memiliki derajat keanggotaan pada tiap-tiap *cluster*. Dengan cara memperbaiki pusat *cluster* dan derajat keanggotaan tiap-tiap titik data secara berulang-ulang, maka didapat lokasi pusat *cluster* optimal. Perulangan ini didasarkan pada minimasi fungsi objektif yang menggambarkan jarak dari titik data yang diberikan ke pusat *cluster* yang terbobot oleh derajat keanggotaan titik data tersebut.

Tujuan penelitian ini adalah untuk meng*cluster*kan data saham harian PT. Astra, Tbk dengan mengimplementasikan metode *Fuzzy C-Means* (FCM). menganalisis *cluster*, kemudian melakukan pelabelan masing-masing *cluster* dan menjadi bahan pertimbangan para investor untuk penanaman modal.

II. TINJAUAN PUSTAKA

Saham

Saham merupakan tanda kepemilikan nilai atau modal badan usaha atau seseorang terhadap sebuah perusahaan (Utami et al., 2019). Seseorang atau badan usaha mempunyai hak atas anggaran dan aset perusahaan, serta dapat hadir saat Rapat Umum Pemegang Saham (RUPS).

Fuzzy C-Means

Fuzzy C-Means (FCM) adalah salah satu metode untuk pengelompokan data dari rumpun ilmu logika fuzzy dan *data mining* berdasarkan nilai keanggotaan. Implementasi FCM didahulukan dengan menentukan nilai kluster, titik data mempunyai nilai keanggotaan yang berbeda pada tiap kluster (Ridho et al., 2020). Dengan melakukan iterasi yang berulang, centroid akan bergerak menuju tempat yang tepat, sehingga terbentuk pengelompokan data. Berikut merupakan langkah-langkah perhitungan FCM:

1. Input data yang akan di *cluster* X, berupa matriks berukuran $n \times m$ (n = jumlah sampel data, m = atribut setiap data). X_{ij} = data sampel ke- i ($i=1,2,\dots,n$), atribut ke- j ($j=1,2,\dots,m$).

2. Tentukan:

- Jumlah *cluster* = c
- Pangkat = w
- Maksimum iterasi = $MaxIter$
- Error terkecil yang diharapkan
- Fungsi objektif awal = $P_0 = 0$
- Iterasi awal = $t = 1$

3. Bangkitkan bilangan random μ_{ik} , $i=1,2,\dots,n$; $k=1,2,\dots,c$; sebagai elemen-elemen matriks partisi awal U.

- Hitung jumlah setiap kolom:

$$Q_i = \sum_{k=1}^c \mu_{ik}$$

- Hitung μ_{ik} ternormalisasi:

$$\mu_{ik} = \frac{\mu_{ik}}{Q_i}$$

4. Hitung pusat *cluster* ke- k : V_{kj} , dengan $k=1,2,\dots,c$; dan $j=1,2,\dots,m$.

$$V_{kj} = \frac{\sum_{i=1}^n (\mu_{ik})^w * X_{ij}}{\sum_{i=1}^n (\mu_{ik})^w}$$

5. Hitung fungsi objektif pada iterasi ke- t , P_t :

$$P_t = \sum_{i=1}^n \sum_{k=1}^c \left(\left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right] (\mu_{ik})^w \right)$$

6. Hitung perubahan matriks partisi:

$$\mu_{ik} = \frac{\left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right]^{-\frac{1}{w-1}}}{\sum_{k=1}^c \left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right]^{-\frac{1}{w-1}}}$$

dengan: $i = 1,2,\dots,n$; dan $k = 1,2,\dots,c$.

7. Cek kondisi berhenti:

- Jika: ($|P_t - P_{t-1}| < x$) atau ($t > MaxIter$) maka berhenti;
- Jika tidak: $t = t+1$, ulangi langkah ke-4.

III. METODOLOGI PENELITIAN

Pendekatan Penelitian

Dalam penelitian ini pendekatan yang digunakan adalah pendekatan kualitatif dan pendekatan kuantitatif. Pada pendekatan kualitatif, dikaji secara literatur tentang metode *Fuzzy C-Means* (FCM) yang diambil dari pustaka. Sedangkan pada pendekatan kuantitatif, menggunakan data sekunder yang diambil dan diolah dari PT Astra Internasional Tbk (ASIIJK).

Metode Pengumpulan Data

Metode pengumpulan data yang digunakan adalah non participant observer, dimana penulis hanya mengamati data yang sudah tersedia tanpa ikut menjadi bagian dari suatu sistem data. Data yang digunakan adalah data saham harian PT. Astra, Tbk periode

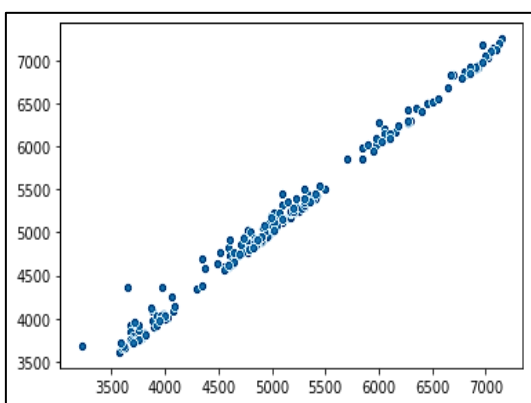
Januari 2020 sampai dengan Oktober 2020.

Metode *Fuzzy C-Means* (FCM) *Fuzzy Clustering* adalah salah satu teknik untuk menentukan *cluster* optimal dalam suatu ruang vektor yang didasarkan pada bentuk normal Euclidean untuk jarak antar vektor. *Fuzzy Clustering* sangat berguna bagi pemodelan fuzzy terutama dalam mengidentifikasi aturan-aturan

fuzzy. Ada beberapa algoritma *clustering* data, salah satu diantaranya adalah *Fuzzy C-Means* (FCM). *Fuzzy C-Means* (FCM) adalah suatu teknik pengclusteran atau pengelompokan data yang mana keberadaan tiap-tiap data dalam satu *cluster* ditentukan oleh fungsi keanggotaan. Teknik ini pertama kali diperkenalkan oleh Jim Bezdek pada tahun 1981 (Kusumadewi, 2004: 83)

IV. HASIL DAN PEMBAHASAN

Dari total 202 data harga saham, akan didistribusikan datanya dan kemudian ditentukan nilai *cluster* berdasarkan distribusi data tersebut. Dapat dilihat data tersebar dan menyatu pada 3 kelompok, dengan ini ditetapkan nilai C adalah 3. Berikut merupakan visualisasi distribusi data.



Gambar 1 Distribusi data

Dengan menggunakan pemrograman bahasa Python, hasil perhitungannya berupa nilai centroid dari setiap *cluster* sebagai berikut:

Tabel 1 Nilai Centroid

Open	High
3835.783	3933.468
6591.937	6648.292
4988.813	5061.854

Nilai - nilai di atas merupakan nilai centroid ketiga *cluster* dengan makna:

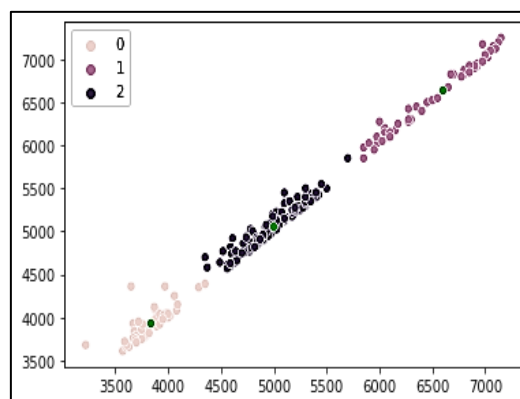
a. *Cluster 0*, nilai centroid *cluster* ini adalah (3835, 3933) yang artinya *cluster 1*

beranggotakan harga saham harian dengan rata - rata harga *open* Rp. 3835 dan rata - rata harga *high* Rp. 3933.

b. *Cluster 1*, nilai centroid *cluster* ini adalah (6591, 6648) yang artinya *cluster 2* beranggotakan harga saham harian dengan rata - rata harga *open* Rp. 6591 dan rata - rata harga *high* Rp. 6648.

c. *Cluster 2*, nilai centroid *cluster* ini adalah (4988, 5061) yang artinya *cluster 0* beranggotakan harga saham harian dengan rata - rata harga *open* Rp. 4988 dan rata - rata harga *high* Rp. 5061.

Dengan demikian distribusi data akan mengelompok berdasarkan nilai centroid masing-masing *cluster*. Berikut merupakan visualisasi *clustering* data harga saham.



Gambar 2 Distribusi data setelah *clustering*

Kemudian, hasil data yang telah diolah dengan pengelompokan 3 *cluster* menggunakan algoritma *Fuzzy C-Means* dapat dilihat lebih jelas dalam bentuk tabel berikut.

	Date	Open	High	Low	Close	Adj Close	Volume	cluster_FCM
0	2020-01-02	6925.0	6925.0	6775.0	6875.0	6213.557129	15008600	2
1	2020-01-03	6925.0	6950.0	6825.0	6950.0	6281.341309	19068800	2
2	2020-01-06	6900.0	6900.0	6750.0	6750.0	6100.583008	22261900	2
3	2020-01-07	6800.0	6825.0	6675.0	6775.0	6123.177734	27963000	2
4	2020-01-08	6700.0	6825.0	6700.0	6775.0	6123.177734	15150400	2
...
197	2020-10-21	5200.0	5250.0	5175.0	5250.0	4927.515137	52567200	0
198	2020-10-22	5200.0	5275.0	5200.0	5250.0	4927.515137	75883000	0
199	2020-10-23	5300.0	5400.0	5250.0	5400.0	5068.300781	46040700	0
200	2020-10-26	5400.0	5450.0	5300.0	5400.0	5068.300781	33055000	0
201	2020-10-27	5300.0	5500.0	5300.0	5425.0	5091.765625	45393900	0

Gambar 3 Data hasil *clustering*

Berdasarkan hasil *clustering*, maka setiap *cluster* dapat dilabelkan dengan kategori tinggi-rendah harga saham dan dianalisis berdasarkan rentang waktu tertentu, sebagai berikut:

- a. *Cluster 0* dilabelkan sebagai '*cluster rendah*' dengan anggota rata-rata harian harga saham *open* dan *high*, masing-masing senilai Rp. 3835 dan Rp. 3933. Anggota *cluster* ini berkelompok pada bulan Januari hingga Maret.
- b. *Cluster 2* dilabelkan sebagai '*cluster sedang*', dengan anggota rata-rata harian harga saham *open* dan *high*, masing-masing senilai Rp. 4988 dan Rp. 5061. Anggota *cluster* ini berkelompok pada bulan Maret hingga Mei.
- c. *Cluster 1* dilabelkan sebagai '*cluster tinggi*' dengan anggota rata-rata harian harga saham *open* dan *high*, masing-masing senilai Rp. 6561 dan Rp. 6648. Anggota *cluster* ini berkelompok pada bulan Mei hingga Desember.

1	Date	Open	High	Low	Close	Adj Close	Volume	cluster_FCM	
43	41	2020-02-28	5850.0	5850.0	5525.0	5525.0	4.993.439.941	100693900	1
44	42	2020-03-02	5700.0	5850.0	5625.0	5700.0	5.151.604.004	63765800	2
45	43	2020-03-03	5850.0	5975.0	5775.0	5900.0	533.236.084	53817300	1
46	44	2020-03-04	5900.0	6025.0	5875.0	5975.0	5.400.145.508	43566800	1
47	45	2020-03-05	6025.0	6050.0	5875.0	6050.0	5.467.929.688	40537300	1
48	46	2020-03-06	5950.0	5950.0	5675.0	5675.0	5.129.008.301	31282500	1
49	47	2020-03-09	5500.0	5500.0	5025.0	5025.0	4.541.544.922	87904100	2
50	48	2020-03-10	5100.0	5325.0	5100.0	5150.0	4.654.519.043	86444800	2
51	49	2020-03-11	5150.0	5175.0	5050.0	5125.0	4.631.923.828	45134700	2
52	50	2020-03-12	5000.0	5025.0	4820.0	4910.0	4.437.609.375	64824200	2
53	51	2020-03-13	4910.0	4910.0	4910.0	4910.0	4.437.609.375	0	2
54	52	2020-03-16	4770.0	4770.0	4530.0	4590.0	4.148.396.484	55655900	2
55	53	2020-03-17	4590.0	4600.0	4270.0	4270.0	3.859.183.594	67166800	2
56	54	2020-03-18	4290.0	4350.0	3980.0	4150.0	3.750.729.004	61972000	0
57	55	2020-03-19	4020.0	4020.0	3860.0	3860.0	3.488.629.639	27066700	0
58	56	2020-03-20	3890.0	3900.0	3590.0	3780.0	341.632.666	141147800	0

Gambar 4 Hasil *clustering* yang tidak sesuai dengan *range* waktu

Pada rentang waktu tertentu, terdapat beberapa data dimana *clustering* berada di *cluster* berbeda. seperti pada data ke-42 yang seharusnya berada di *cluster 1* namun masuk ke *cluster 2*. Perbedaan *cluster* ini terjadi karena harga *open* dan *high* pada data ke-42 berada pada *range cluster 2*. Hal demikian juga terjadi pada data ke-47 hingga data ke-52. Adapun

penyebabnya yakni karena sifat saham adalah fluktuatif. Harga saham ditentukan oleh perusahaan disesuaikan dengan kondisi dan prospek perusahaan. Sebagai tambahan, isu perusahaan juga dapat memengaruhi besar kecilnya nilai harga saham.

KESIMPULAN

Berikut ini adalah beberapa kesimpulan yang dapat diambil berdasarkan pembahasan yang telah dilakukan sebelumnya :

1. *Fuzzy C-Means* (FCM) merupakan metode *clustering* data yang digolongkan sebagai metode pengklasifikasian yang bersifat unsupervised (tanpa arahan),
2. Pengalokasian ulang data ke dalam masing-masing *cluster*, metode C-Means menggunakan metode pengalokasian yang bersifat tegas (hard) sedangkan untuk metode FCM memanfaatkan teori fuzzy,
3. Pengelompokkan harga saham harian PT. Astra, Tbk., dengan metode FCM menghasilkan 3 *cluster* yaitu '*cluster rendah*' dengan anggota pada bulan Januari - Maret rata-rata harian harga saham *open* Rp. 3835 dan *high* Rp. 3933, '*cluster sedang*', anggota pada bulan Maret - Mei dengan rata-rata harian harga saham *open* Rp. 4988 dan *high* Rp. 5061, dan '*cluster tinggi*' pada bulan Mei - Desember dengan anggota rata-rata harian harga saham *open* Rp. 6561 dan *high* Rp. 6648.
4. Terdapat 7 data yang tidak berkelompok sesuai dengan *range* bulan.

Saran yang dapat diberikan yaitu menggunakan dataset dengan jumlah lebih besar dan menggunakan metode lainnya sebagai perbandingan metode terbaik sehingga dapat menjadi acuan untuk penelitian selanjutnya.

DAFTAR PUSTAKA

- [1] Febrian, Muhammad Fitri. "Penerapan *Fuzzy C-Means* Dan *Fuzzy Time Series* Untuk Analisis Nilai Saham BEI Tahun 2015-2020."
- [2] Handoyo, S., & Efendi, A. (2019). Generating of Fuzzy Rule Bases with Gaussian Parameters Optimized via Fuzzy C-Mean and Ordinary Least Square. *Blue Eyes Intelligence Engineering & Sciences Publication*
- [3] Hasibuan, R. A., Rizki, S., W., & Perdana, H., (2021). Perbandingan Metode Fuzzy Clustering Means Dan Single Linkage Pada Pengelompokan Saham LQ45. *Buletin Ilmiah Math. Stat. dan Terapannya (Bimaster)*, 10(3), 361 - 368
- [4] Purnawati, I., & Werastuti, D. (2013). Faktor Fundamental Ekonomi Makro Terhadap Harga Saham LQ45. *Jurnal Keuangan dan Perbankan*, 17(2), 211 - 219.
- [5] Rahmatika, L., Suparti, S., & Safitri, D. (2015). Analisis Kelompok Dengan Algoritma *Fuzzy C-Means* Dan Gustafson Kessel Clustering Pada Indeks LQ45. *Jurnal Gaussian*, 4(3), 543-551.
- [6] Ridho, M., Priadi, R., Trianasari, N., Si, S., Stat, M., Prodi, S., Bisnis, M., Ekonomi, F., & Telkom, U. (2020). Analisis Pengelompokan Perusahaan Sektor Keuangan Utama Berdasarkan Nilai Harga Saham Menggunakan *Fuzzy C-Means Clustering Grouping Analysis of Sector Companies Primary Finance Based on Stock Price Value Using Fuzzy C-Means Clustering*. 7(2), 2095–2104.
- [7] Sari, S. K., Tarno, T., & Safitri, D. (2017). Pemilihan Input Model Regression Adaptive Neuro Fuzzy Inference System (Ranfis) Untuk Kajian Data IHSIG. *Jurnal Gaussian*, 6(3), 449-458.
- [8] Sumarauw, S. J. A., & Subanar. (2007). *Fuzzy C-Means Clustering Model Data Mining For Recognizing Stock Data Sampling Pattern*. IJCCS, 1(2), 115 - 122
- [9] Utami, M. P., Saepudin, D., & Atiqi, A. R. (2019). Pengaruh Teknik Clustering Harga Saham Dalam Manajemen Portofolio Effects of Clustering Stock Price Techniques in Portfolio Management. *E-Proceeding of Engineering*, 6(1), 2491–2509.
- [10] Utamil, M., P., Saepudin, D., & Rohmawati, A., A., (2019). Pengaruh Teknik Clustering Harga Saham Dalam Manajemen Portofolio. *e-Proceeding of Engineering*, 6(1), 2491
- [11] Sudirman. (2015). *Pasar modal dan manajemen portofolio.Gorontalo: Sultan Amai Press*.
- [12] Sumarauw, Sylvia. "Fuzzy C-Means Clustering untuk Pengenalan Pola Studi kasus Data Saham." *Jurnal Axioma: Jurnal Matematika dan Pembelajaran* 7.2 (2022): 97-106.