



Penerapan Data Mining Tentang Perbandingan Pemahaman *Non-Fungible Token* (NFT) Menggunakan *K-Means*

Fari Adidan^{1*}, Perani Rosyani²

¹Fakultas Teknik, Program Studi Ilmu Komputer, Universitas Pamulang, Tangerang Selatan, Indonesia

Email: ^{1*}didan.adidan@gmail.com, ²dosen00837@unpam.ac.id

(* : coresponding author)

Abstrak – Arsip digital memiliki karakteristik yang mudah disalin atau diperbanyak sehingga dengan hal tersebut sangat sulit atau bahkan tidak memungkinkan untuk menjual suatu karya seni berbentuk digital seperti halnya menjual karya seni berbentuk konvensional. Namun kini, dengan perkembangan teknologi *blockchain* dan mata uang kripto yang berkembang pesat menjadikan hal yang disebutkan tadi menjadi mungkin dengan cara mengubah karya seni tersebut menjadi sebuah NFT atau *Non-Fungible Token*. Sebuah karya seni digital yang berbentuk NFT ini bisa menjadi peluang besar untuk dijadikan keuntungan bagi para pembuat karya seni tersebut. Namun demikian, masih banyak masyarakat yang hanya ingin mengikuti trend saja tanpa mengetahui pemahaman secara mendalam apa itu NFT dan bagaimana cara kerjanya dalam dunia kearsipan. Oleh sebab itu, penelitian ini bertujuan untuk mengetahui seberapa banyak perbandingan pada masyarakat terhadap pemahaman tentang NFT tersebut, sehingga dengan mengetahui hasil perbandingan nya diharapkan dapat memberikan pemahaman secara pengelompokkan kepada masyarakat yang belum banyak memahami konsep NFT.

Kata Kunci: Arsip Digital, *Non-Fungible Token*, *Clustering*, Perbandingan Pemahaman

Abstract – Digital archives have the characteristic of being easily copied or reproduced, so that it is very difficult or even impossible to sell a digital work of art as it is to sell a conventional form of art. But now, with the development of blockchain technology and rapidly growing cryptocurrencies, this has become possible by turning the work of art into an NFT or Non-Fungible Token. A digital artwork in the form of an NFT can be a great opportunity to make profits for the creators of the artwork. However, there are still many people who just want to follow the trend without having an in-depth understanding of what NFT is and how it works in the world of archives. Therefore, this study aims to find out how many comparisons there are in the community with regard to the understanding of the NFT, so knowing the results of the comparison is expected to provide a grouping understanding of the community who do not understand the concept of NFT much.

Keywords: Digital Archives, Non-Fungible Tokens, Clustering, Comparison of Understanding

1. PENDAHULUAN

Non-Fungible Token (NFT) merupakan sebuah arsip digital pada bidang kriptografi. NFT adalah jenis aset digital baru yang memungkinkan transaksi lebih efisien dan aman. NFT dapat mewakili koleksi hal-hal atau sesuatu seperti uang, sekuritas, catatan objek atau data (Houser & Holden, 2022). *Non-Fungible Token* atau NFT sempat populer beberapa waktu lalu di Indonesia oleh seorang kolektor yang mempunyai nama asli Sultan Gustaf Al Ghozali atau lebih dikenal dalam sosial media nya dengan nama Ghozali Everyday. Berkat kepopuleran nya, banyak masyarakat yang mulai mengikuti *trend* dengan membuat dan menjual NFT. Namun terdapat masalah utama yang timbul akibat *trend* ini, yaitu kurangnya pemahaman pada masyarakat dalam mengunggah sesuatu objek menjadi sebuah NFT. Pasalnya, bukan sebuah karya seni, uang ataupun aset berharga lainnya, melainkan hal-hal aneh yang tidak seharusnya dijual dan diunggah untuk dijadikan arsip digital berupa NFT, misalnya seperti Kartu Tanda Penduduk (KTP), foto berupa makanan, minuman, atau barang-barang di sekitar, bahkan sampai ada yang menjual foto bayi/anak-anak mereka. Hal tersebut sangat melenceng jauh dengan kegunaan NFT yang sebenarnya. Tidak seharusnya data-data pribadi kita yang bersifat privasi di unggah sembarangan apalagi diperjual-belikan dalam bentuk NFT.

Oleh karena itu, solusi untuk permasalahan ini yaitu memberikan sebuah pemahaman kepada masyarakat yang lebih spesifik, seperti pengertian dasar NFT, bagaimana cara kerjanya, serta bagaimana cara yang benar dalam menggunakan data-data kita yang akan dijadikan sebuah aset digital berupa NFT.



Penelitian yang dilakukan oleh Mhd Gading Sadewo pada tahun 2018 dengan judul PENERAPAN ALGORITMA CLUSTERING DALAM MENGELOMPOKKAN BANYAKNYA DESA/KELURAHAN MENURUT UPAYA ANTISIPASI/MITIGASI BENCANA ALAM MENURUT PROVINSI DENGAN K-MEANS membahas tentang banyaknya data pada desa/kelurahan dalam rangka upaya antisipasi/mitigasi bencana alam menurut kelompok provinsi. Terdapat 4 variabel yang digunakan yaitu Sistem Peringatan Dini Bencana Alam, Sistem Peringatan Dini Tsunami, Perlengkapan Keselamatan, dan Jalur Evakuasi. Metode yang digunakan dalam penelitian ini adalah menggunakan algoritma k-means clustering. Penelitian ini memiliki kesimpulan bahwa 3 provinsi tingkat tinggi yaitu Jawa Barat, Jawa Tengah, Jawa Timur, 9 provinsi tingkat sedang, dan 22 provinsi lainnya termasuk tingkat rendah (Sadewo et al., 2018).

Penelitian yang dilakukan oleh Sukma Sindi pada tahun 2020 dengan judul ANALISIS ALGORITMA K-MEDOIDS CLUSTERING DALAM PENGELOMPOKAN PENYEBARAN COVID-19 DI INDONESIA mengangkat permasalahan tentang data mining terhadap pengelompokan kasus penyebaran covid-19 di Indonesia. Variabel yang digunakan dalam penelitian ini ada tiga, yaitu kasus positif, kasus sembuh, dan kasus meninggal. Penelitian ini menggunakan metode clustering yaitu algoritma k-medoids. Kesimpulan pada penelitian ini adalah bahwa dari 34 record data diperoleh 1 record pada cluster pertama, 2 record pada cluster kedua, 31 record pada cluster ketiga (Sindi et al., 2020).

Penelitian yang dilakukan oleh Yudi Irawan pada tahun 2019 yang berjudul PENERAPAN DATA MINING UNTUK EVALUASI DATA PENJUALAN MENGGUNAKAN METODE CLUSTERING DAN ALGORITMA HIRARKI DIVISIVE memiliki permasalahan tentang penerapan data mining terhadap data penjualan mesin dan bahan-bahan digital printing perusahaan Media World Pekanbaru tahun 2018. Penelitian ini menggunakan beberapa atribut yaitu No. Item, No. Faktur, Tgl Faktur, Kuantitas, Unit Barang, dan Nama Pelanggan. Metode clustering yang digunakan dalam penelitian ini adalah algoritma hirarki divisive. Kesimpulan pada penelitian ini adalah menghasilkan pengetahuan berupa customer loyal yang ada di perusahaan Media World Pekanbaru dan setelah dievaluasi dengan menggunakan algoritma hirarki divisive serta pengolahan data menggunakan software rapid miner maka ditemukan bahwa customer loyal berada pada cluster 3 dengan range 117-358 yaitu Mater Q, Cash, Istana Print, Sahabat Adv, Jasa Reklame, Brilian Adv, Kreasi Adv, Family Print, Dian Print, Banner Teddy, Galaxi, Multi Baru, Print Art, Master Print, Zoom Reklame, Media World Jambi, Image Print dan WSN Adv (Irawan, 2019).

2. METODE

Metode yang dilakukan pada penelitian ini adalah dengan menggunakan salah satu metode algoritma clustering yaitu k-means. Kemudian data akan diolah menggunakan rumus yang ada pada metode k-means yaitu euclidean distance (ED).

2.1 Metode Pengumpulan Data

a. Studi Literatur

Dalam penelitian ini, studi literatur dilakukan dengan cara mengumpulkan referensi-referensi dari jurnal, artikel, dan internet yang berkaitan dengan metode-metode dan studi kasus.

b. Kuesioner

Kuesioner dilakukan dalam penelitian ini digunakan untuk mendapatkan data-data yang dibutuhkan dari hasil survei yang berhubungan dengan studi kasus terkait.

2.2 Data Mining

2.2.1 Seleksi Data

Pada tahap ini, data akan di seleksi dengan cara melakukan penghapusan data yang tidak digunakan dalam penelitian. Karena data tersebut akan memperlambat proses pengujian data. Kemudian data-data tersebut akan dipilih sesuai ketentuan penelitian agar dapat diolah untuk pengujian data selanjutnya.

Pada penelitian ini, data-data akan dipilih berdasarkan jawaban responden terbanyak yang diambil dari 2 kota/kabupaten yaitu Bogor dan Tangerang dengan total sebanyak 100 data responden. Kemudian data akan dipilih masing-masing 50 data untuk setiap wilayah kota/kabupaten tersebut sehingga didapat 2 data atribut untuk *dataset* nya.

2.2.2 Transformasi Data

Pada tahap ini, transformasi data dilakukan agar data tersebut dapat lebih mudah diolah. *Dataset* yang sudah diseleksi sebelumnya akan diubah ke bentuk *dataset* yang lebih sederhana sesuai dengan kebutuhan pengujian data yang akan dilakukan yaitu mengambil total skor nya saja dengan 2 atribut sebagai X dan Y. Berikut ini adalah hasil transformasi data :

Tabel 1. Transformasi Data

Data ke-	X	Y
1	42	39
2	18	21
3	39	36
4	30	21
5	21	30
6	33	18
7	36	33
8	42	21
9	21	30
10	33	30
11	27	18
12	30	24
13	45	30
14	24	33
15	30	12
16	21	30
17	12	27
18	39	30
19	36	27
20	30	33
21	36	30
22	30	27
23	27	24
24	30	36
25	27	39
26	27	18

27	33	24
28	30	27
29	24	15
30	33	39
31	21	30
32	30	24
33	18	33
34	42	18
35	30	9
36	27	33
37	24	27
38	36	27
39	30	27
40	24	36
41	42	21
42	45	30
43	24	15
44	21	33
45	33	45
46	42	15
47	27	33
48	27	24
49	30	42
50	45	39

2.2.3 K-Means Clustering

Dataset yang sebelumnya sudah di transformasi, selanjutnya akan dihitung menggunakan metode *k-means* dengan rumus *Euclidean Distance* (ED) untuk menghitung proses *clustering*.

Berikut adalah langkah-langkah pengujian data menggunakan metode *k-means clustering* :

- a. Menentukan jumlah kluster yang ingin dibuat. $k=2$ (Karena jumlah *cluster* berdasarkan 2 wilayah kota/kabupaten yang telah dipilih yaitu Kabupaten Bogor dan Kota Tangerang).
- b. Menentukan pusat kluster (*centroid*) awal pada *dataset*, peneliti mengambil data ke-1 dan data ke-2 sebagai *centroid* awal :

Tabel 2. Centroid Awal

Data ke-	X	Y
1	42	39
2	18	21

- c. Menentukan kelompok kluster dengan menghitung jarak setiap data terhadap pusat kluster (*centroid*) awal yaitu pada data ke-1 dan data ke-2. Untuk menentukan kelompok kluster, ambil hasil perhitungan dengan jarak terdekat (nilai minimum). Data uji akan dihitung menggunakan *Euclidean Distance* dengan rumus sebagai berikut :

$$D(x, y) = \sqrt{(x1 - x2)^2 + (y1 - y2)^2}$$

Berikut ini adalah hasil pengujian data yang sudah dikelompokkan berdasarkan kluster :

Tabel 3. Hasil Pengelompokkan *Cluster*

Data ke-	X	Y	Cluster
1	42	39	1
2	18	21	2
3	39	36	1
4	30	21	2
5	21	30	2
6	33	18	2
7	36	33	1
8	42	21	1
9	21	30	2
10	33	30	1
11	27	18	2
12	30	24	2
13	45	30	1
14	24	33	2
15	30	12	2
16	21	30	2
17	12	27	2
18	39	30	1
19	36	27	1
20	30	33	1
21	36	30	1
22	30	27	2
23	27	24	2
24	30	36	1
25	27	39	1
26	27	18	2
27	33	24	2



28	30	27	2
29	24	15	2
30	33	39	1
31	21	30	2
32	30	24	2
33	18	33	2
34	42	18	1
35	30	9	2
36	27	33	2
37	24	27	2
38	36	27	1
39	30	27	2
40	24	36	2
41	42	21	1
42	45	30	1
43	24	15	2
44	21	33	2
45	33	45	1
46	42	15	1
47	27	33	2
48	27	24	2
49	30	42	1
50	45	39	1

d. Setelah mendapatkan hasil pengelompokkan cluster, langkah selanjutnya yaitu meng-update nilai pusat klaster (*centroid*) baru yang dapat dihitung dengan cara mencari nilai rata-rata data yang menjadi anggota tiap klaster sebagai berikut :

1. Untuk klaster 1 (C1) terdapat 21 data yaitu data ke-1, data ke-3, data ke-7, data ke-8, data ke-10, data ke-13, data ke-18, data ke-19, data ke-20, data ke-21, data ke-24, data ke-25, data ke-30, data ke-34, data ke-38, data ke-41, data ke-42, data ke-45, data ke-46, data ke-49, dan data ke-50, sehingga :

$$X_{\text{baru}} = \frac{42+39+36+42+33+45+39+36+30+36+30+27+33+42+36+42+45+33+42+30+45}{21} = 37.28$$

$$Y_{\text{baru}} = \frac{39+36+33+21+30+30+30+27+33+30+36+39+39+18+27+21+30+45+15+42+39}{21} = 31.42$$

2. Untuk klaster 2 (C2) terdapat 29 data yaitu data ke-2, data ke-4, data ke-5, data ke-6, data ke-9, data ke-11, data ke-12, data ke-14, data ke-15, data ke-16, data ke-17, data ke-22, data ke-23, data ke-26, data ke-27, data ke-28, data ke-29, data ke-31, data ke-32, data ke-33, data ke-35, data ke-36, data ke-37, data ke-39, data ke-40, data ke-43, data ke-44, data ke-47 dan data ke-48, sehingga :

Berikut adalah hasil pusat kluster baru (centroid) yang telah diperbarui :

Tabel 4. Centroid Baru

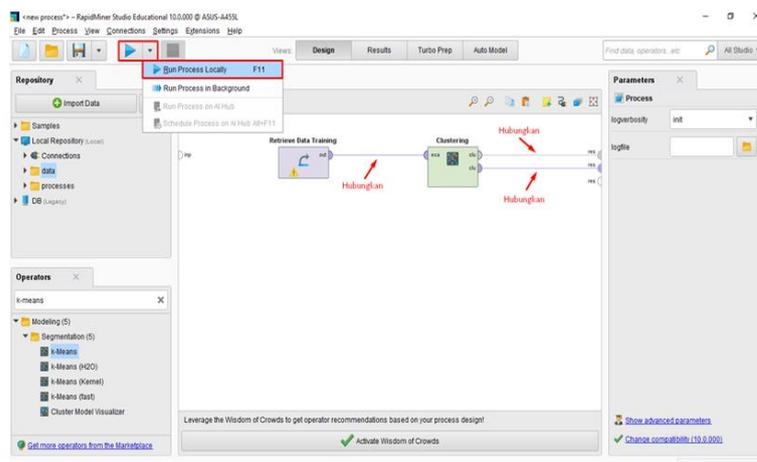
Data ke-	X	Y
1	37.28	31.42
2	25.55	24.93

e. Jika posisi *cluster* tidak ada yang berubah, maka proses *clustering* dapat dihentikan.

3. ANALISA DAN PEMBAHASAN

3.1 Implementasi Sistem

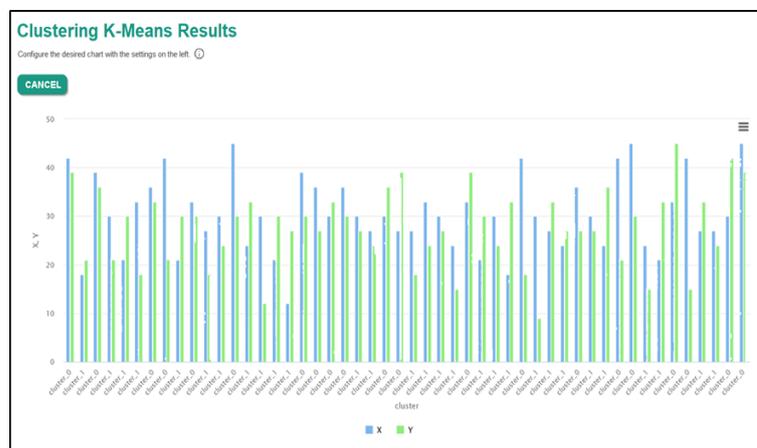
Melakukan proses perhitungan dan hasil implementasi k-means dari dataset menggunakan aplikasi RapidMiner Studio 10.0. Langkah-langkah nya dimulai dari input dataset, menentukan operator clustering, mengatur parameter, lalu menghubungkan setiap komponen dan memulai proses eksekusi perhitungan sebagai berikut :



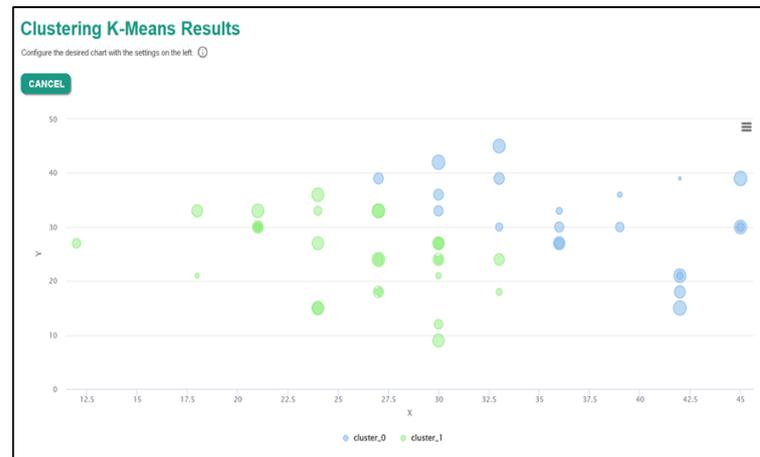
Gambar 1. Proses Perhitungan K-Means

3.2 Hasil Pengujian

Setelah melakukan proses perhitungan menggunakan *software* RapidMiner, didapat hasil pengujian berupa grafik dan plot cluster sebagai berikut :



Gambar 2. Hasil K-Means Clustering Berupa Bar Chart



Gambar 3. Hasil *K-Means Clustering* Berupa *Plot Cluster*

4. KESIMPULAN

Berikut adalah kesimpulan yang didapat dari hasil implementasi dan pengujian data, didapat kesimpulan sebagai berikut :

- a. Cara yang tepat untuk mengunggah data-data pribadi untuk dijadikan sebuah arsip digital berupa NFT yaitu dilarang menggunakan data yang berhubungan dengan privasi individu maupun kelompok masyarakat. Seperti foto wajah seseorang, kartu identitas, maupun dokumen yang bersifat rahasia. Hal tersebut dapat melanggar UU ITE dalam membuat suatu karya cipta seperti NFT ini.
- b. Menurut hasil pengujian data clustering dengan total 100 data responden yang telah dihitung menggunakan metode k-means dengan rumus euclidean distance, didapatkan hasil pengelompokan clustering yaitu sebanyak 21 data pada klaster 1 yaitu untuk wilayah Kabupaten Bogor, dan sebanyak 29 data pada klaster 2 yaitu untuk wilayah Kota Tangerang.
- c. Dengan hasil tersebut, maka dapat ditarik kesimpulan bahwa masyarakat yang berada di kelompok wilayah Kota Tangerang lebih memahami tentang Non-Fungible Token (NFT) dibandingkan dengan masyarakat yang berada di kelompok wilayah Kabupaten Bogor.

REFERENCES

- [1] K. A. Houser and J. T. Holden, "Navigating the Non-Fungible Token Kimberly A. Houser* & John T. Holden**," *SSRN Electron. J.*, pp. 1–36, 2022.
- [2] M. G. Sadewo, A. P. Windarto, and A. Wanto, "Penerapan Algoritma Clustering Dalam Mengelompokkan Banyaknya Desa/Kelurahan Menurut Upaya Antisipasi/ Mitigasi Bencana Alam Menurut Provinsi Dengan K-Means," *KOMIK (Konferensi Nas. Teknol. Inf. dan Komputer)*, vol. 2, no. 1, pp. 311–319, 2018, doi: 10.30865/komik.v2i1.943.
- [3] S. Sindi *et al.*, "Analisis algoritma k-medoids clustering dalam pengelompokan penyebaran covid-19 di indonesia," vol. 4, no. 1, pp. 166–173, 2020.
- [4] Y. Irawan, "Penerapan data mining untuk evaluasi data penjualan menggunakan metode clustering dan algoritma hirarki divisive," *Jtiulm*, vol. 3, no. 1, pp. 13–20, 2019.
- [5] M. U. Noor, "NFT (NON-FUNGIBLE TOKEN): MASA DEPAN ARSIP DIGITAL? ATAU HANYA SEKEDAR BUBBLE?," *Pustakaloka*, vol. 13 No.2, pp. 1–12, 2021.
- [6] S. Sundari, S. N. Faiza, and L. Rahma, "Business Analysis of Selling and Buying Non-Fungible Tokens (NFT) at Marketplace OpenSea according to The Perspective of Islamic Law," *El-Qist J. Islam. Econ. Bus.*, vol. 12, no. 1, pp. 1–16, 2022, doi: 10.15642/elqist.2022.12.1.1-16.
- [7] T. Alfina and B. Santosa, "Analisa Perbandingan Metode Hierarchical Clustering, K-Means dan Gabungan Keduanya dalam Membentuk Cluster Data (Studi Kasus : Problem Kerja Praktek Jurusan Teknik Industri ITS)," *Anal. PerbandinganMetode Hierarchical Clust. K-means dan Gabungan Keduanya dalam Clust. Data*, vol. 1, no. 1, pp. 1–5, 2012.