

BAB II

LANDASAN TEORI

2.1 *Non-Fungible Token*

Non-Fungible Token (NFT) merupakan *token* dalam bentuk aset digital yang berbasis *blockchain* untuk memetakan hak kepemilikan dari aset digital tersebut secara aman. Beberapa contoh dari aset digital dalam bentuk NFT adalah karya seni berupa lukisan, musik, hingga *game* [18]. NFT dapat diperjualbelikan dengan menggunakan mata uang *cryptocurrencies*, seperti Ethereum. Beberapa *marketplace* populer yang memperdagangkan NFT, yaitu OpenSea, Mintable, dan Rarible. Salah satu *collection* NFT yang terjual dengan harga yang cukup tinggi adalah karya dari Beeple yang berjudul “*Everydays:the First 5000 Days*” yang terjual seharga 69 juta USD pada bulan Maret 2021 [2].

Kekurangan dari NFT adalah tidak ada pengukuran pasti mengenai *investment performance* maupun prediksi harga dari *collection* NFT. NFT dapat dibeli dengan didasarkan pada keyakinan dan kepercayaan dari *value* dari *collection* NFT itu sendiri [19]. Selain itu, baik *creator* atau *owner* dan *collector* dari *collection* NFT tidak memiliki kontrol terhadap karya digital tersebut. Hal ini dalam artian bahwa sebagai *creator* ataupun *collector* tidak bisa mengendalikan orang-orang lain yang membicarakan, tertarik, ataupun menggunakan NFT tersebut. Namun, jika dilihat dari sisi positifnya, *creator* ataupun *collector* cenderung akan senang ketika orang-orang banyak membicarakan mengenai NFT yang dimilikinya, menggunakan, hingga menyebarluaskan hal yang berkaitan dengan *collection* NFT tersebut. Aktivitas tersebutlah yang mampu memberikan pengaruh yang besar pada kepentingan dan popularitas dari *collection* NFT tersebut [20].

Kelebihan dari NFT adalah dapat dijadikan salah satu peluang bagi *creator* untuk menjangkau lebih banyak penggemar dari karya yang diciptakannya. Partisipasi *creator* pada dunia NFT dinilai baik untuk mampu meningkatkan kesadaran orang-orang mengenai seni, terutama dalam bentuk karya digital [4]. Pemasaran karya seni digital dengan membuat NFT memberikan keuntungan lain di luar dari hasil penjualan NFT, yaitu dengan *royalty*. *Creator* akan memperoleh

royalty dari setiap kali NFT yang dibuat oleh *creator* tersebut terjual ke *owner* barunya[5].

Terdapat beberapa istilah pada NFT yang umum digunakan, antara lain: [21]

- a. *Creator*
Creator merupakan orang yang membuat dan memasarkan *collection* maupun *item* NFT di *marketplace* NFT.
- b. *Owner*
Owner merupakan orang yang memiliki *item* NFT di suatu *marketplace* NFT.
- c. *Collection*
Collection merupakan kumpulan dari *item* NFT yang dibuat oleh *creator* yang memiliki *design* yang sama namun variasinya berbeda. Perbedaan dari *item* NFT ini menunjukkan tingkat kelangkaan (*rare*) dari *item* NFT tersebut.
- d. *Royalty*
Royalty merupakan penghasilan dalam bentuk persentase yang diperoleh *creator* setiap kali NFT yang dibuat oleh *creator* tersebut terjual ke *owner* barunya.
- e. *Burn*
Burn merupakan proses untuk menghancurkan atau menghilangkan suatu NFT dari suatu sirkulasi dengan memindahkan *item* tersebut ke *wallet address* yang tidak dapat diakses.
- f. *Delist*
Delist merupakan proses menghilangkan *item* NFT dari transaksi di suatu *marketplace*.
- g. *Ethereum*
Ethereum merupakan suatu sistem *blockchain* dengan memiliki fungsionalitas *smart contract* sebagai *platform* utama bagi projek-projek NFT.
- h. *Floor price*
Floor price merupakan harga *item* NFT yang paling rendah pada satu *collection* NFT.

i. *Fractional ownership*

Fractional ownership merupakan pembagian kepemilikan *item* NFT, dimana penjual menjual *item* dalam bentuk persentase, sehingga pembeli dapat membeli sebanyak dengan yang mereka mampu.

j. *Metadata*

Metadata merupakan kumpulan data yang mendefinisikan kepemilikan dari *item* NFT dan pembeda antara satu *item* NFT dengan *item* NFT lainnya.

k. *MetaMask*

MetaMask merupakan *Ethereum wallet gateway* yang paling banyak digunakan menghubungkan *wallet* ke *marketplace* NFT.

2.2 Framework, Algoritma Clustering, dan Metode Evaluasi

2.2.1 Framework

2.2.1.1 Data Mining

Data mining merupakan proses yang dilakukan untuk menggali informasi tersembunyi pada kumpulan data yang belum diketahui sebelumnya dengan proses manual [22]. *Data mining* bertujuan untuk mengungkap pola-pola pada data yang digunakan sebagai jalan keluar dari permasalahan yang dihadapi dan menghasilkan prediksi berdasarkan pola-pola yang ditemukan.

Data mining secara sederhana dibedakan menjadi dua tipe, yaitu *predictive* dan *descriptive*. Secara lebih detail, *data mining* dikelompokkan ke dalam beberapa teknik, yaitu: [23]

1. *Clustering*

Clustering merupakan salah satu teknik *data mining* untuk melakukan klasifikasi pada data. Tujuan dari dilakukannya *clustering* adalah untuk menemukan jarak ataupun kesamaan dari suatu data.

2. *Classification*

Classification merupakan teknik *data mining* untuk melakukan klasifikasi dengan mengungkap pola atau model yang dapat mendeskripsikan dan membedakan suatu kelas data.

3. *Regression*

Regression merupakan teknik data mining untuk melakukan analisa pada data guna menemukan hubungan antar variabel. Tujuan dari penggunaan *regression* adalah untuk melakukan evaluasi terhadap kemungkinan pada variabel tertentu berdasarkan variabel lainnya.

4. *Association rules*

Association rules merupakan teknik *data mining* untuk menemukan pola pada sekumpulan data. Selain itu, *association rules* digunakan untuk menemukan hubungan antara dua ataupun lebih *variable*.

5. *Outer detection*

Outer detection merupakan teknik *data mining* untuk melakukan pengamatan pada *variable* tertentu pada data yang tidak berhubungan atau tidak berkorelasi dengan pola yang akan digunakan pada prediksi.

6. *Sequential patterns*

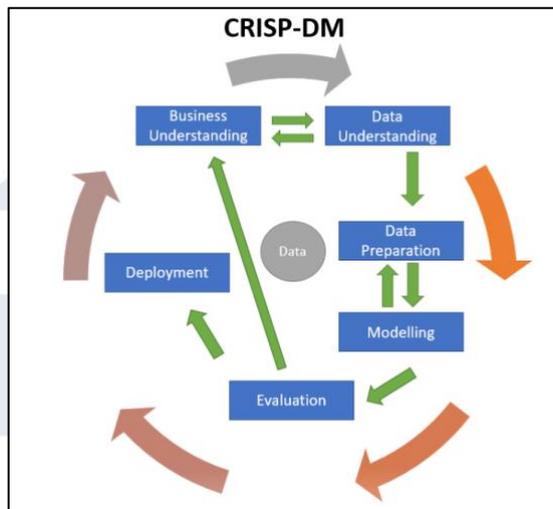
Sequential patterns merupakan teknik *data mining* untuk mengungkap pola pada data yang memiliki kemiripan pola dalam jangka waktu tertentu. *Sequential patterns* banyak digunakan pada data transaksi.

7. *Predictions*

Predictions merupakan teknik *data mining* yang dimana pada penggunaannya juga memanfaatkan teknik *data mining* lainnya untuk membantu dalam melakukan prediksi suatu kejadian di masa depan. Beberapa teknik *data mining* yang digunakan bersama dengan *prediction*, seperti *sequential patterns*, *clustering*, *classification*, dan lain sebagainya.

2.2.1.2 CRISP-DM

Cross Industry Standard Process for Data Mining (CRISP-DM) merupakan *framework* model yang digunakan pada *data mining*. CRISP-DM terdiri dari 6 tahapan, yaitu *business understanding*, *data understanding*, *data preparation*, *modelling*, *evaluation*, dan *deployment* [24].



Gambar 2.1 Framework CRISP-DM

Sumber: Towards Data Science (2021)

Berikut merupakan 6 tahapan dari *framework* CRISP-DM, yaitu: [25]

1. *Business understanding*

Business understanding merupakan tahapan untuk menentukan tujuan yang ingin dicapai oleh bisnis secara jangka panjang dan menerapkan tujuan bisnis dan diterjemahkan menjadi tujuan dari dilakukannya proses *data mining*.

2. *Data understanding*

Data understanding merupakan tahapan pemahaman terhadap data yang akan digunakan, mulai dari mengumpulkan data, mendeskripsikan data, melakukan evaluasi pada data, dan pemilihan atribut yang akan digunakan.

3. *Data preparation*

Data preparation merupakan tahapan yang melakukan pemilihan atribut yang akan digunakan, menghapus data *null* atau data *redundant*, dan melakukan transformasi pada data.

4. *Modelling*

Modelling merupakan tahapan pemilihan dan pembentukan model yang sesuai dengan data yang dimiliki untuk memperoleh hasil yang optimal [26].

5. *Evaluation*

Evaluation merupakan tahapan untuk memastikan bahwa setiap tahapan tidak ada yang terlewat mulai dari *business understanding* hingga *modelling*. Selain itu, tahapan ini juga harus memastikan bahwa *model* yang terbentuk telah sesuai dengan standar dari algoritma pemodelan yang dibangun [27].

6. *Deployment*

Deployment merupakan tahapan untuk melakukan penyebaran informasi yang diperoleh dalam bentuk laporan dan dengan analisa yang dapat dengan mudah dimengerti oleh pihak terkait [26].

2.2.2 *Algoritma Clustering*

2.2.2.1 *K-Means*

K-Means merupakan algoritma yang digunakan untuk teknik klasifikasi dengan melakukan pengelompokan pada data yang sama menjadi satu *cluster*. Data yang berada pada *cluster* yang sama memiliki karakteristik, yaitu tingkat variasinya kecil [28]. K-Means merupakan salah satu dari beberapa algoritma *unsupervised learning*. K-Means tergolong kedalam algoritma *unsupervised learning* karena penggunaannya untuk mengungkap pola baru dari atribut pada data yang belum ditemukan sebelumnya dan dibagi ke dalam beberapa *cluster* dengan partisi [29].

Dalam membentuk suatu kelas *cluster*, algoritma K-Means hanya mendukung untuk pengelolaan data yang bersifat numerik atau dapat pula dikatakan untuk pengelolaan data kuantitatif. Pada dasarnya, algoritma K-Means ini merupakan algoritma yang bersifat iterative dengan tahapan sebagai berikut: [11]

1. Pada *dataset* yang ada, tentukan nilai k sebagai jumlah *cluster* yang ingin dibentuk
2. Menentukan *centroid* atau titik pusat dari *cluster* secara *random*
3. Melakukan perhitungan jarak masing-masing titik data ke *centroid* dengan mengukur jarak terdekat terhadap *centroid* tersebut

4. Melakukan pembentukan *cluster* dengan mengelompokkan objek berdasarkan jarak titik data dengan *centroid* terdekat
5. Melakukan iterasi pada tahapan ke-3 dan ke-4 hingga menemukan *centroid* yang optimal.

2.2.2.2 K-Medoids

K-Medoids merupakan varian dan pengembangan dari algoritma K-Means yang mampu mengatasi kelemahan yang dimiliki oleh K-Means. Model *cluster* dengan menggunakan algoritma K-Medoids ini lebih kebal terhadap *outlier* dibandingkan dengan algoritma K-Means. Hal ini dikarenakan algoritma K-Means sangat dipengaruhi oleh nilai *mean*. Pada K-Medoids, pengelompokan dilakukan dengan menjadikan *medoids* sebagai titik aktual pada pembentukan *cluster*. *Medoids* merupakan suatu titik atau objek pada suatu *cluster* yang memiliki jarak terhadap titik lain yang minimum [11].

Penentuan *medoids* sebagai titik pusat awal pada *cluster* merupakan bentuk representasi dari *cluster* itu sendiri. Oleh karena itu, jumlah objek *medoids* yang ditentukan sejumlah dengan jumlah *cluster* yang ingin dibentuk. Tahapan pemodelan dengan algoritma K-Medoids, yaitu: [30]

1. Pada *dataset*, tentukan nilai k sebagai jumlah *cluster* yang ingin dibentuk
2. Melakukan pemetaan pada titik data pada masing-masing *cluster* dengan melakukan perhitungan jarak menggunakan *Euclidian Distance*.
3. Pada masing-masing *cluster*, tentukan titik *medoids* sebagai representasi dari *cluster* tersebut
4. Melakukan perhitungan jarak pada setiap titik data terhadap *medoids* pada tiap *cluster*
5. Melakukan perhitungan nilai simpangan yang merupakan nilai jarak baru dikurangi dengan jarak yang lama. Apabila nilai simpangan < 0 , objek akan ditukar dengan titik data pada *cluster* yang dilanjutkan dengan pembentukan *medoids* baru dengan kumpulan k objek yang baru.

6. Melakukan iterasi pada tahapan ke-3 hingga ke-5 hingga tidak adanya perubahan pada *medoids*

2.2.2.3 X-Means

X-Means juga merupakan salah satu algoritma pengembangan dari algoritma K-Means. Penggunaan dari algoritma ini ditujukan untuk mengatasi kelemahan dari algoritma K-Means yang mengharuskan melakukan penentuan nilai k secara manual dan biaya komputasi yang tergolong tinggi. Oleh karena itu, X-Means hadir untuk memberikan solusi dalam bentuk pemberian estimasi dari nilai k atau jumlah *cluster* optimal yang terbentuk dengan lebih efisien dan tidak memerlukan biaya komputasi yang tinggi namun mampu memberikan proses komputasi yang efisien pula [12].

Secara sederhana, cara kerja dari algoritma X-Means adalah menentukan dan memberi keputusan terhadap pembagian lebih lanjut terhadap *centroid* dengan melakukan pencarian sela antara jumlah *cluster* dengan tempat *cluster* [31]. Tahapan pembentukan *cluster* dengan algoritma X-Means ditempuh dengan langkah berikut ini: [32]

1. Menentukan kemungkinan nilai k *min* dan k *max* berdasarkan *dataset* yang digunakan
2. Gunakan nilai k *min* sebagai nilai k yang digunakan pada pembentukan model K-Means (umumnya digunakan k *min* bernilai 2)
3. Lakukan pembentukan model K-Means
4. Terhadap model *cluster* yang terbentuk kemudian akan dipecah menjadi dua *child cluster*
5. Lakukan pengecekan kriteria terhadap masing-masing *cluster* untuk menentukan apakah *cluster* tersebut akan dipecah atau tidak.
6. Dari hasil yang diperoleh pada tahapan ke-5, secara otomatis jumlah *cluster* yang terbentuk akan diperbaharui
7. Melakukan iterasi pada tahapan ke-3 hingga ke-6 pada setiap *cluster*

2.2.3 Metode Evaluasi

2.2.3.1 Davies Bouldin Index

Davies Bouldin Index (DBI) merupakan salah satu dari beberapa parameter untuk melakukan evaluasi pada model *cluster*. Evaluasi pada model *cluster* dengan menggunakan *Davies Bouldin Index* didasarkan pada dua poin penilaian, yaitu jumlah kerapatan data pada *cluster* yang terbentuk terhadap *centroid* dan jarak antar *centroid* terhadap *cluster* yang bersangkutan. Penentuan model terbaik didasarkan oleh semakin kecil nilai DBI yang dihasilkan oleh model tersebut [29].

2.3 RapidMiner, Table Capture, Microsoft Excel, dan Tableau

2.3.1 RapidMiner



Gambar 2.2 Logo RapidMiner

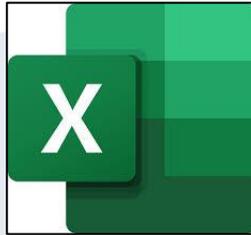
Sumber: RapidMiner (2022)

RapidMiner merupakan *software* yang dapat digunakan untuk membantu proses dalam melakukan pengolahan pada data. Fitur-fitur yang terdapat pada RapidMiner digunakan untuk melakukan *data preparation*, *data mining*, *machine learning*, visualisasi hingga analisa prediktif [33].

2.3.2 Table Capture

Table Capture merupakan salah satu *extension* yang disediakan oleh Google Chrome. Table Capture digunakan untuk membantu dalam melakukan *scraping data*, yaitu dengan menyalin tabel yang terdapat di *file HTML* atau *website* tertentu untuk kemudian dikonversikan ke dalam bentuk *spreadsheet*.

2.3.3 Microsoft Excel



Gambar 2.3 Logo Microsoft Excel

Sumber: Google (2019)

Microsoft Excel merupakan salah satu dari *tools* yang paling banyak digunakan dalam melakukan pengolahan data dalam bentuk *spreadsheet*. Penggunaan dari Microsoft Excel sendiri tidak hanya ditujukan untuk *programmer* saja, namun kemudahan dari penggunaan *tools* ini menjadikan *tools* ini mampu untuk digunakan oleh semua kalangan [34].

2.3.4 Tableau



Gambar 2.4 Logo Tableau

Sumber: Google (2021)

Tableau merupakan salah satu *tools* dalam visualisasi data yang mampu mengolah data dalam jumlah yang sangat besar. Pertama kali diluncurkan pada tahun 2003, Tableau mampu menghasilkan visualisasi yang canggih dengan hanya menggunakan fungsi *drag and drop*. Tableau diperlengkapi dengan berbagai jenis visualisasi, baik *bar chart*, *plots*, *heatmap* dan lain sebagainya [35].

2.4 Penelitian Terdahulu

Tabel 2.1 Penelitian Terdahulu

Nama Jurnal	Judul Artikel	Penulis	Metode	Hasil
<i>Telecommunication Computing Electronics</i>	<i>Cluster-based Water Level Patterns Detection</i>	Friska Natalia Ferdinand, Yustinus	K-Medoids, DBSCAN, dan X-Means	<i>Clustering</i> berdasarkan variabel <i>water level</i> dengan

Nama Jurnal	Judul Artikel	Penulis	Metode	Hasil
<i>and Control</i> (Telkomnika) Vol.17, No.3, 2019 [13]		Soelistio, Ferry Vincenttius Ferdinand, I Made Murwantara		K-Medoids dan X-Means menghasilkan 3 cluster dengan persebaran data yang lebih merata dibandingkan dengan DBSCAN
Jurnal Pendidikan Teknologi dan Kejuruan Vol. 18, No. 2, 2021 [14]	Analisis Pemilihan <i>Cluster</i> Optimal dalam Segmentasi Pelanggan Toko Retail	Santi Ika Murpratiwi, I Gusti Agung Indrawan, Arik Aranta	K-Means, K-Medoids, dan X- Means	Perbandingan algoritma berdasarkan nilai DBI menunjukkan algoritma K- Medoids lebih baik digunakan dibandingkan dua algoritma lainnya dengan nilai DBI sebesar 0,540778
<i>International Journal of Information Technology and Computer Science</i> Vol.12, No.5, 2020 [15]	<i>Data Mining Methods for Detecting the Most Significant Factors Affecting Students' Performance</i>	Mohammed Abdullah Al-Hagery, Maryam Abdullah Alzaid, Tahani Soud Alharbi, Moody Abdulrahman Alhanaya	K-Means dan X- Means	Kedua algoritma menunjukkan hasil performa yang sama, yaitu 0.138
Jurnal Ilmiah Rekayasa dan Manajemen Sistem Informasi, Vol. 5, No. 1, 2019 [16]	Perbandingan Algoritma K- Means dan K- Medoids untuk Pengelompokan Data Transaksi Bongkar Muat	Insanul Kamila, Ulya Khairunnisa , Mustakim	K-Means dan K- Medoids	K-Means memiliki nilai DBI yang lebih rendah dibandingkan K-Medoids dengan nilai 0.112 banding 0.119

Nama Jurnal	Judul Artikel	Penulis	Metode	Hasil
	di Provinsi Riau			
Jurnal ICT: <i>Information Communication & Technology</i> , Vol.20, No.2, 2021 [17]	Perbandingan Algoritma K-Means, X-Means, dan K-Medoids Untuk Klasterisasi Awak Kabin Lion Air	Ahmad Jurnaidi Wahidin, Dana Indra, Sensuse	K-Means, K-Medoids, dan X-Means	K-Means lebih cocok untuk diterapkan dengan nilai DBI sebesar 0,792. Setelah itu, disusul oleh X-Means dengan DBI 0,812 dan K-Medoids memiliki DBI sebesar 1,700

Berdasarkan Tabel 2.1 dan mengacu pada penelitian [13] menghasilkan *output* dalam bentuk visualisasi data berupa *dashboard*. Penelitian ini dilakukan dengan membandingkan tiga algoritma *clustering*, yaitu K-Means, K-Medoids, dan X-Means dengan nilai DBI sebagai pembanding sebagaimana dilakukan pada penelitian [14] dan [17]. Namun, pembeda dengan penelitian [14] dan [17] adalah pada objek penelitiannya, yaitu *collection* NFT. Pembeda lebih lanjut dari penelitian [15] dan [16] adalah jumlah dan kombinasi algoritma *clustering* yang digunakan. Selain itu, penelitian ini juga menggunakan RapidMiner sebagai *tools* untuk melakukan pengelolaan pada data, sebagaimana digunakan pada penelitian [16].