Prediksi Nilai *Non-Fungible Token* denganAlgoritma *Decision Tree*

R. FADHLAN AUFFAR[1*], FAHMI ARIF [2], ALIF ULFA[3],

AFIFAHInstitut Teknologi Nasional Bandung
Email penulis: auffarfadhlan@mhs.itenas.ac.id

Received 29 01 2023 | *Revised* 05 02 2023 | *Accepted* 05 02 2023

ABSTRAK

Non-Fungible Token (NFT) adalah aset digital yang disimpan dan diedarkan pada jaringan blockchain. Terdapat dua jenis NFT yaitu collectible item dan utility item. Nilai transaksi NFT pada tahun 2021 sangat tinggi, beberapa NFT dapat terjual dengan harga ratusan ribu dolar. NFT yang dapat terjual tinggi menandakan NFT bisa menjadi peluang bisnis digital baru untuk bisa mendapatkan keuntungan dengan cara memiliki NFT bernilai tinggi. Hal tersebut menandakan untuk dapat memiliki NFT bernilai tinggi perlu informasi yang dapat membantu untuk menaksir prospek NFT yang ada. Berdasarkan fenomena tersebut dibutuhkan sebuah model yang dapat menaksir nilai prospek NFT. Salah satu metode yang dapat digunakan untuk menanggulangi hal tersebut adalah supervised learning khususnya klasifikasi. Metodologi yang digunakan adalah CRISP-DM. Algoritma yang digunakan untuk membuat model klasifikasi adalah decision tree. Algoritma tersebut memiliki accuracy score sebesar 0,75, precision 0,64, recall 0,45, dan f1-score 0,52.

Kata kunci: NFT, supervised learning, prediksi, klasifikasi, CRISP-DM

ABSTRACT

Non-Fungible Token (NFT) is digital asset that are stored and circulated on blockchain network. There are two types of NFT, namely collectible items and utility items. The value of NFT transactions in 2021 is very high, some NFTs canbe sold for hundred thousand dollars. This indicates that NFTs can become a new digital business opportunity to be able benefit by owning high value NFTs. Thisimplies that in order to have a high value NFT, information is needed that canhelp to assess the prospect for existing NFTs. Based on this, a model is neededthat can estimate the value of NFT prospect. One method that can be used toovercome this is supervised learning, especially classification. The methodology used is CRISP-DM. The algorithm used to create the classification model is decision tree. The algorithm has an accuracy score of 0,75, precision of 0,64, recall of 0,45, and f1-score of 0,52.

Keywords: NFT, supervised learning, prediction, classification, CRISP-DM

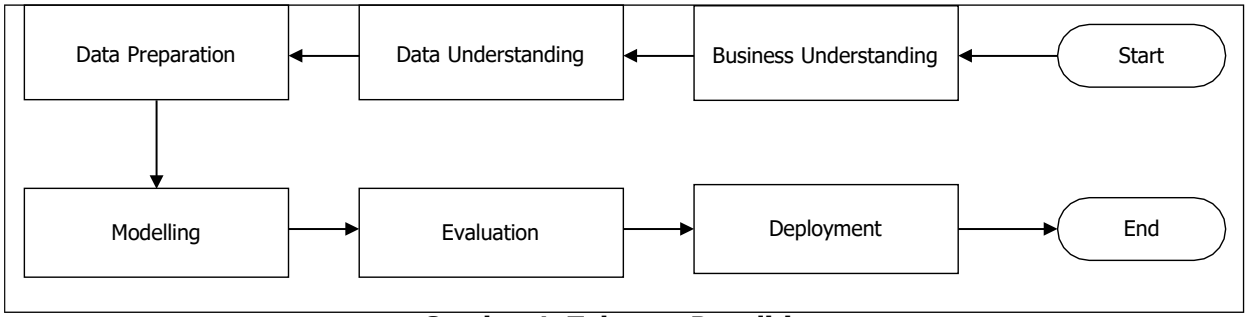
1. PENDAHULUAN

NFT merupakan singkatan dari *Non-Fungible Token*, NFT merupakan tanda kepemilikan dari aset digital seperti gambar, video, musik, dan *virtual world item* (Dowling, 2022). Ekosistem dari NFT mengalami ledakan pasar tahun 2020 (NonFungible, 2021). Banyak NFT terjual dengan harga tinggi. Nilai transaksi NFT tahun 2021 sebesar USD 17 miliar, serta kenaikan harga rata-rata dari NFT dari tahun 2020 ke tahun 2021 naik sebesar 1542% (*NonFungible*, 2022). Banyak faktor yang mempengaruhi NFT. Sentimen pasar, social media, transaction *fee*, infrastruktur *blockchain,* dan kebijakan dapat mempengaruhi nilai NFT (Ante, 2022). Nilai harga dari NFT tidak dapat divaluasi secara langsung berdasarkan sistem ekonomi, melainkan dari sentimen pasar (Kapoor et al., 2022). Kenaikan dari harga rata-rata NFT mengindikasikan bahwa NFT bisa menjadi peluang bisnis digital baru. Nilai harga jual NFT dari awal NFT diluncurkan sampai nilai tersebut direspon oleh pasar mengalami perubahan nilai yang signifikan. Untuk bisa mendapatkan keuntungan dari NFT perlu memiliki NFT yang bernilai tinggi, sehingga diperlukan informasi mengenai nilai prospek dari NFT yang ada.

Nilai prospek NFT yang tinggi dapat diketahui dengan cara membuat model prediksi yang dapat menaksir serta mengklasifikasi prospek NFT yang ada berdasarkan faktor yang mempengaruhi nilai tersebut. Salah satu metode yang dapat digunakan untuk menanggulangi persoalan tersebut adalah dengan bantuan *machine learning* khususnya supervised learning. *Supervised learning* dapat digunakan apabila terdapat input berupa *independent variable* (X) dan output berupa *dependent variable* (Y), yang dimana terdapat penggunaan algoritma untuk mempelajari fungsi-fungsi persamaan dari variabel tersebut (Brownlee, 2016). Penggunaan *supervised learning* dapat membantu untuk membuat model prediksi karena terdapat algoritma yang dapat digunakan untuk menaksir serta mengklasifikasi nilai dari prospek NFT, selain dari itu terdapat bantuan mesin untuk melakukan komputasi dimensi data yang besar serta jenis data yang beragam.

2. METODOLOGI PENELITIAN

Metodologi yang digunakan dalam penelitian ini adalah *Cross Industry Standard Process for Data Mining* (CRISP-DM). CRISP-DM merupakan metode dari data mining yang menyediakan langkah-langkah komprehensif (Shearer, 2000). Terdapat 6 tahap proses yang dilakukan, tahapan tersebut adalah *business understanding, data understanding*, *data preparation*, *modelling, evaluation,* dan *deployment* (Provost & Fawcett, 2013).



Gambar 1. Tahapan Penelitian

Proses pertama yang telah dilakukan adalah *business understanding*. Pada proses ini terdapat beberapa tahap yang telah dilakukan yaitu adalah menganalisis bisnis dan pasar NFT serta merumuskan permasalahan. Publik mulai mencermati NFT disebabkan karena nilai transaksinya yang tinggi (Nadini et al., 2021). Minat untuk mengoleksi NFT naik pada tahun 2021 (Mekacher et al., 2022). Hal tersebut menandakan pasar NFT dapat menjadi peluang

bisnis baru untuk bisa mendapatkan keuntungan. Untuk bisa mendapatkan keuntungan perlu memiliki NFT dengan nilai tinggi. Nilai NFT berubah secara signifikan ketika NFT diluncurkan sampai pasar merespon nilai tersebut. Berdasarkan fenomena tersebut dibutuhkan model yang dapat membantu dalam menaksir serta mengklasifikasi nilai dari prospek NFT yang ada. Untuk menyelesaikan persoalan tersebut dapat dibantu dengan *supervised learning*. *Supervised learning* dapat menemukan antara input (*independent variable*) dengan *output* (*dependent variable*) (Soofi & Awan, 2017). Sehingga dengan penggunaan *supervised learning* diharapkan dapat mengetahui nilai prospek NFT berdasarkan faktor yang mempengaruhinya.

Proses kedua yang telah dilakukan adalah data *understanding.* Proses ini bertujuan untuk mengidentifikasi kebutuhan data yang diperlukan untuk menyelesaikan permasalahan, mengumpulkan, mendeskripsikan, dan mengeksplorasi data yang sudah ditentukan. Untuk mengatasi persoalan dalam memprediksi nilai prospek NFT membutuhkan *feature* data yang dapat merepresentasikan nilai NFT. *Feature* yang dapat merepresentasikan hal tersebut adalah nilai harga dari NFT itu sendiri. Sentimen pasar mempengaruhi nilai NFT, sehingga perlu *feature* data yang dapat merepresentasikan hal tersebut. Data yang dibutuhkan juga perlu menggambarkan perubahan nilai dari NFT tersebut pertama kali diluncurkan sampai pasar merespon nilai tersebut. Data yang terkumpul didapat dari situs Kaggle. Data tersebut tersedia dari hasil *web scraping* nftswohroom.com oleh Arjan De Haan pada tanggal 28 Oktober 2021. Data memiliki dimensi 4189 baris yang merepresentasikan NFT dan 15 kolom yang merepresentasikan *feature*. Data tersebut digunakan karena memiliki *feature* yang dapat merepresentasikan variabel yang dibutuhkan dalam penelitian ini.

Proses eksplorasi data dilakukan pada media *jupyter notebook* dengan bahasa pemrograman *python.* Dalam proses eksplorasi data langkah yang dilakukan adalah pemilihan *feature,* penyaringan data, penentuan *dependent* dan *independent variable*, eksplorasi karakteristik data, pemeriksaan *outlier,* mencari data *error*, dan menghitung nilai statistika data. Pemilihan *feature* didasari berdasarkan deskripsi data. *Feature* yang digunakan dapat dilihat pada Tabel 1.

Tabel 1. Pemilihan *Feature*

Feature	**Jenis**	**Deskripsi**
Creator	*Categorical*	Pembuat NFT
Price	*Numerical*	Harga NFT
Type	*Categorical*	Tipe NFT
Likes	*Numerical*	Jumlah *like* yang didapatkan
NSFW	*Categorical*	Label untuk menunjukan konten NFT (*Not Safe For Work*)
Tokens	*Numerical*	Jumlah NFT yang dibuat
Year	*Categorical*	Tahun pembuatan NFT

Feature creator, type, likes, dan NSFW dipilih karena dapat menggambarkan sentimen pasar. *Feature year* dipilih karena dapat merepresentasikan waktu NFT dibuat. *Feature price* digunakan karena dapat menggambarkan nilai NFT, sedangkan *feature tokens* dipilih karena dapat mengetahui supply NFT. Setelah feature terpilih proses selanjutnya adalah penyaringan data, penyaringan data dilakukan untuk mengetahui perubahan nilai NFT dari tahun NFT dibuat sampai dengan waktu dari hasil *web scraping*. Penyaringan data dilakukan kepada NFT yang dibuat pada tahun 2020. Hal tersebut mengakibatkan dimensi baris data berkurang menjadi 2368 baris*. Feature price* dijadikan sebagai *dependent variable* karena sesuai dengan tujuan penelitian ini, sedangkan *feature* lainnya dijadikan *independent variable. Feature* yang berjenis *numerical data* dilakukan pengecekan data menggunakan histogram. Berdasarkan hasil histogram *feature* mengalami kemiringan (*skewness*). Kemiringan pada data menyebabkan model klasifikasi cenderung salah saat mengklasifikasi

kelas minoritas (Larasati et al., 2019). Setelah itu *feature* tersebut diperiksa *outlier* menggunakan visualisasi *boxplot*. Berdasarkan hasil *boxplot* beberapa *feature* memiliki *outlier* yang banyak.

Eksplorasi data dilanjutkan untuk *feature* berjenis *categorical data* dengan cara mencari jumlah label untuk setiap *feature* kemudian dilanjut dengan mencari frekuensi untuk setiap label tersebut. *Feature creator* memiliki jumlah label sebanyak 340, yang menggambarkan nama-nama dari pembuat NFT. Feature type memiliki 3 label *yaitu photo, gif,* dan *video. Feature* NSFW memiliki 2 label yaitu *true* dan *false*. Apabila label menunjukan *true* maka konten NFT tersebut berisikan konten yang tidak cocok dilihat pada saat bekerja. Proses selanjutnya yang dilakukan adalah pemeriksaan data *error.* Maksud data *error* disini adalah apakah ada data yang bernilai kosong atau tidak. Setelah itu ada perhitungan nilai statistika contohnya seperti persentil dan rata-rata.

Proses ketiga yang telah dilakukan adalah *data preparation*. Kegunaan proses ini adalah untuk mendapatkan data yang berkualitas untuk diproses. Proses yang telah dilakukan adalah data *cleaning* serta *data transformation*. Proses data cleaning dilakukan pada data yang *error,* namun dikarenakan pada tahap eksplorasi data tidak ditemukan data yang *error* maka data tidak perlu dibersihkan. Proses yang dilakukan pada data transformation adalah *transformasi numerical data*, *encoding categorical data*, *discretization*, serta pembagian data. Proses *transformasi numerical data* dengan *cara log transformation*. Metode *logarithmic transformation* sederhana namun sangat berguna meningkatkan validitas data (Choi et al., 2019). Rumus dari metode tersebut dapat dilihat sebagai berikut.

$$Y_i' = \log (Y_i + 1) \quad (1)$$

Ket:
Y_i' = Data ke-i sudah ditransformasi
Y_i = Data ke-i sebelum ditransformasi

Encoding categorical data dilakukan dengan cara *one hot encoding. Feature creator* akan di-*encoding* berdasarkan 10 label dengan frekuensi terbesar. *Feature type* akan di-*encoding* untuk label photo dan gif untuk menghindari *dummy variable trap*. *Encoding feature* NSFW dengan cara mengubah label *true* menjadi 1 dan *false* menjadi 0. Setelah proses *encoding* selesai selanjutnya adalah *discretization*. Proses tersebut bertujuan untuk melihat nilai tinggi atau rendahnya dari *feature price*. Proses tersebut dilakukan dengan cara membagi kelas dari *feature price* menjadi 2 kelas. Kelas 0 (rendah) memiliki rentang nilai dari 0 sampai dengan 5 sedangkan kelas 1 (tinggi) memiliki rentang nilai lebih besar dari 5 sampai dengan 15. Nilai 5 diambil dari persentil 75 sedangkan 15 diambil dari nilai maksimal *feature price* setelah ditransformasi. Setelah proses encoding selesai, selanjutnya adalah pembagian data. Data dibagi menjadi 2 jenis data yaitu train data dan test data.

Proses keempat yang telah dilakukan adalah *modelling*. Proses ini dilakukan pada media *jupyter notebook* dengan bahasa pemrograman *python*. Algoritma yang digunakan pada proses pembuatan model adalah *Decision Tree* (DT). Algoritma ini digunakan karena efisien dan mudah dipahami selain itu algoritma ini merupakan metode yang kuat serta umum digunakan untuk berbagai macam bidang (Charbuty & Abdulazeez, 2021). Proses kelima yang telah dilakukan adalah *evaluation*. Proses ini bertujuan untuk mengevaluasi model prediksi menggunakan *confusion matrix*, selain itu perhitungan indikator performansi yaitu *accuracy score, precision, recall,* dan *F1-score*. Rumus dari indikator performansi dapat dilihat sebagai berikut.

$$\text{Accuracy score} = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (3)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (4)$$

$$\text{F1-Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (5)$$

Ket:
TP = True positive
FP = False positive
TN = True negative
FN = False negative

Proses keenam adalah *deployment*. Proses ini bertujuan untuk memaparkan hasil model dari algoritma terpilih kemudian mengimplementasikan model terpilih menggunakan 10 data NFT baru yang diambil pada tanggal 20 November 2022. Data tersebut dapat dilihat pada Tabel 2.

Tabel 2. Data Baru NFT

No	*Creator*	*Type*	*Likes*	NSFW	*Token*
1	tblings-art	*Photo*	1	FALSE	4
2	badsexy	*Photo*	3	FALSE	2
3	akida	*Photo*	0	FALSE	3
4	pittcn	*Photo*	1	FALSE	3
5	pittcn	*Photo*	1	FALSE	3
6	driptorchpress	*Photo*	0	FALSE	3
7	mikiad.visuals	*Photo*	0	FALSE	3
8	spottydogg	*Photo*	0	FALSE	3
9	kingdannys	*Gif*	0	FALSE	3
10	rowell	*Gif*	0	FALSE	1

3. HASIL DAN ANALISIS

Hasil dari pemilihan feature dan penyaringan data dapat dilihat pada Tabel 3. Dimensi data menjadi 2368 baris dan 6 kolom.

Tabel 3. Hasil Pemilihan Feature **dan Penyaringan Data**

No	*creator*	*price*	*type*	*likes*	NSFW	*tokens*
1	kristyglas	50	PHOTO	2	FALSE	30
2	juliakponsford	500	VIDEO	0	FALSE	1
3	badsexy	10	PHOTO	0	TRUE	2
…	…	…	…	…	…	…
2366	rubenalexander	50	PHOTO	0	FALSE	3
2367	elgeko	99	GIF	0	FALSE	7
2368	elgeko	700	PHOTO	0	FALSE	1

Hasil pengecekan sebaran data untuk *feature price, likes, dan tokens* dengan menggunakan histogram mendapatkan hasil bahwa sebaran data dari *feature* tersebut memiliki sifat kemiringan (*skewness*). Hasil dari pengecekan outlier dari *feature price, like, dan tokens* menggunakan *boxplot* memperlihatkan bahwa setiap *feature* memiliki *outlier*. *Feature* yang memiliki *outlier* terbanyak adalah *feature tokens.* Hasil pengecekan jumlah label yang dimiliki *feature* berjenis categorical data dapat dilihat pada Tabel 4.

Tabel 4. Hasil Pengecekan Label

Feature	Jumlah Label
Creator	340
Type	3
NSFW	2

Label yang memiliki frekuensi terbanyak untuk *feature creator* adalah richardfyaters dengan frekuensi sebesar 97. Frekuensi label photo memiliki frekuensi sebesar 1661, gif sebesar 487, dan video sebesar 220 pada *feature type*, sedangkan Label *true* memiliki frekuensi sebesar 74 sedangkan sisanya label *false* pada feature NSFW. Hasil pemeriksaan data kosong menghasilkan bahwa untuk setiap *feature* tidak memiliki data kosong.

Hasil transformasi *feature* berjenis *numerical* data menggunakan metode *log transformation* mengakibatkan nilai dari *feature* tersebut berubah. Proses transformasi tersebut menyebabkan sebaran data yang dicek kembali menggunakan histogram berubah yang memperlihatkan bahwa sifat kemiringan (*skewness)* berkurang, selain itu pengecekan outlier yang dilakukan kembali menggunakan boxplot memperlihatkan bahwa *outlier* berkurang. Hasil dari *encoding feature* berjenis categorical data mengakibatkan *feature* bertambah karena proses encoding dilakukan untuk 10 label frekuensi terbesar pada *feature creator* serta untuk label photo dan *gif pada feature type.*

Proses *discretization* mengakibatkan *feature price* menjadi memiliki kelas 0 berjumlah 1640 sedangkan kelas 1 berjumlah 728. Proses pembagian data dilakukan untuk membuat X *train*, X *test*, Y *train*, dan Y *test*. Hasil modelling berdasarkan algoritma *decision tree* dapat dilihat pada Tabel 5.

Tabel 5. Hasil *Modelling*

Algoritma	Jumlah Prediksi Nilai Rendah (Kelas 0)	Jumlah Prediksi Nilai Tinggi (Kelas 1)
Decision tree	372	102

Hasil dari proses *evaluation* dari algoritma *decision tree* dapat dilihat pada Tabel 6.

Tabel 6. Hasil *Evaluation*

Indikator	Hasil
Accuracy score	0,75
Precision	0,64
Recall	0,45
F1-*score*	0,52

Hasil proses *deployment* terdiri dari rekapitulasi model serta implementasi model. Rekapitulasi dari model prediksi berdasarkan algoritma *decision tree* dapat dilihat sebagai berikut:

1. Jika $X_2 \leq 0{,}9$ dan $X_7 \leq 0{,}5$ dan $X_3 \leq 0{,}5$ dan $X_6 \leq 0{,}5$ maka NFT bernilai tinggi.
2. Jika $X_2 \leq 0{,}9$ dan $X_7 \leq 0{,}5$ dan $X_3 \leq 0{,}5$ dan $X_6 > 0{,}5$ maka NFT bernilai rendah.

3. Jika $X_2 \leq 0{,}9$ dan $X_7 \leq 0{,}5$ dan $X_3 > 0{,}5$ maka NFT bernilai rendah.
4. Jika $X_2 \leq 0{,}9$ dan $X_7 > 0{,}5$ maka NFT bernilai rendah.
5. Jika $X_2 > 0{,}9$ dan $X_{10} \leq 0{,}5$ dan $X_{12} \leq 0{,}5$ maka NFT bernilai rendah.
6. Jika $X_2 > 0{,}9$ dan $X_{10} \leq 0{,}5$ dan $X_{12} > 0{,}5$ dan $X_2 \leq 1{,}5$ maka NFT bernilai tinggi.
7. Jika $X_2 > 0{,}9$ dan $X_{10} \leq 0{,}5$ dan $X_{12} > 0{,}5$ dan $X_2 > 1{,}5$ maka NFT bernilai rendah.
8. Jika $X_2 > 0{,}9$ dan $X_{10} > 0{,}5$ dan $X_{14} \leq 0{,}5$ maka NFT bernilai tinggi.
9. Jika $X_2 > 0{,}9$ dan $X_{10} > 0{,}5$ dan $X_{14} > 0{,}5$ dan $X_2 \leq 1{,}59$ maka NFT bernilai tinggi.
10. Jika $X_2 > 0{,}9$ dan $X_{10} > 0{,}5$ dan $X_{14} > 0{,}5$ dan $X_2 > 1{,}59$ maka NFT bernilai rendah.

Keterangan:
X_0 = likes
X_1 = NSFW
X_2 = tokens
X_3 = *Creator*_richardfyates
X_4 = *Creator* _elgeko
X_5 = *Creator*_doze
X_6 =*Creator*_elenasteem
X_7 = *Creator* _rektdoteth
X_8 = *Creator*_artxmike
X_9 = *Creator* _deeanndmathews
X_{10} = *Creator* _juliakponsford
X_{11} = *Creator* _hairofmedusa
X_{12} = *Creator* _barbarabezina
X_{13} = *type*_PHOTO
X_{14} = *type*_VIDEO

Hasil implementasi 10 data NFT baru terhadap algoritma *decision tree* menghasilkan bahwa 9 dari NFT tersebut memiliki nilai prospek yang rendah sedangkan sisanya bernilai tinggi. Hasil tersebut dapat dilihat pada Tabel 9.

Tabel 9. Hasil Implementasi

No	Hasil Kelas Prediksi	Arti
1	0	NFT bernilai rendah
2	0	NFT bernilai rendah
3	0	NFT bernilai rendah
4	0	NFT bernilai rendah
5	0	NFT bernilai rendah
6	0	NFT bernilai rendah
7	0	NFT bernilai rendah
8	0	NFT bernilai rendah
9	0	NFT bernilai rendah
10	1	NFT bernilai tinggi

4. KESIMPULAN

Algoritma *decision tree* dapat digunakan untuk memprediksi nilai prospek NFT. Algoritma tersebut memiliki hasil *accuracy score* sebesar 0,75*, precisio*n sebesar 0,64, *recall* sebesar 0,45, dan f1-*score* sebesar 0,52*. Accuracy score* tersebut menandakan bahwa hasil prediksi berdasarkan model tersebut memiliki 75% kemungkinan bahwa hasil prediksi tersebut benar. Hasil implementasi algoritma tersebut berdasarkan data baru dari 10 NFT menghasilkan

bahwa 9 NFT bernilai rendah dan 1 NFT bernilai tinggi. Model yang telah dibuat tidak dapat digeneralisir menggunakan data pelatihan yang berbeda.

DAFTAR PUSTAKA

Ante, L. (2022). The Non-Fungible Token (NFT) Market and Its Relationship with Bitcoin and Ethereum. FinTech, 1(3), 216–224. https://doi.org/10.3390/fintech1030017

Brownlee, J. (2016). Master Machine Learning Algorithms Discover How They Work and Implement Them from Scratch. http://MachineLearningMastery.com

Charbuty, B., & Abdulazeez, A. (2021). Classification Based on Decision Tree Algorithm for Machine Learning. Journal of Applied Science and Technology Trends, 2(01), 20–28. https://doi.org/10.38094/jastt20165

Choi, J. H., Kim, J., Won, J., & Min, O. (2019). Modelling Chlorophyll-a Concentration using Deep Neural Networks considering Extreme Data Imbalance and Skewness. International Conference on Advanced Communication Technology, ICACT, 2019-February, 631–634. https://doi.org/10.23919/ICACT.2019.8702027

Dowling, M. (2022). Fertile LAND: Pricing non-fungible tokens. Finance Research Letters, 44. https://doi.org/10.1016/j.frl.2021.102096

Kapoor, A., Guhathakurta, D., Mathur, M., Yadav, R., Gupta, M., & Kumaraguru, P. (2022). TweetBoost: Influence of Social Media on NFT Valuation. http://arxiv.org/abs/2201.08373

Larasati, A., Hajji, A. M., & Dwiastuti, A. (2019). The relationship between data skewness and accuracy of Artificial Neural Network predictive model. IOP Conference Series: Materials Science and Engineering, 523(1). https://doi.org/10.1088/1757-899X/523/1/012070

Mekacher, A., Bracci, A., Nadini, M., Martino, M., Alessandretti, L., Aiello, L. M., & Baronchelli, A. (2022). Heterogeneous rarity patterns drive price dynamics in NFT collections. Scientific Reports, 12(1). https://doi.org/10.1038/s41598-022-17922-5

Nadini, M., Alessandretti, L., di Giacinto, F., Martino, M., Aiello, L. M., & Baronchelli, A. (2021). Mapping the NFT revolution: market trends, trade networks, and visual features. Scientific Reports, 11(1). https://doi.org/10.1038/s41598-021-00053-8

NonFungible. (2021). Yearly NFT Market Report 2020 - Everything You Need to Know About the NFT Ecosystem.

NonFungible. (2022). Yearly NFT Market Report 2021 - How NFT Affects the World.

Provost, F., & Fawcett, T. (2013). Data Science for Business.

Shearer, C. (2000). The CRISP-DM Model: The New Blueprint for Data Mining. Journal of Data Warehousing, 5, 13–22.

Soofi, A. A., & Awan, A. (2017). Classification Techniques in Machine Learning: Applications and Issues. Journal of Basic & Applied Sciences, 13, 459–465.