

ANALISIS DATA TENTANG KLASIFIKASI IKAN AIR TAWAR DIBATAM DENGAN MENGGUNAKAN DECISION TREE

(Data analysis on the classification of freshwater in Batam using rapidminer)

¹Almira Yolanda Rahman, ²Cristina Manullang, ³Angela Huang, ⁴Winnie Yuliarni, ⁵Joozu Adzikra Suiyoubi

^{1,2,3,4,5}Program Studi Manajemen Rekayasa, Institut Teknologi Batam

E-mail: 2212008@student.iteba.ac.id, 2212046@student.ac.id, 2212021@student.iteba.ac.id, 2212027@student.iteba.ac.id,
2012018@student.iteba.ac.id

KEYWORDS:

Classification, Forecasting, Freshwater Fish, Freshwater Fish Production, Decision Trees, Batam

ABSTRACT

Freshwater fish cultivation in Batam has great potential to improve the local economy and meet the needs of national and international markets. However, business actors face challenges in optimizing production and sales due to large and complex data. This research aims to overcome this problem by applying data mining techniques using classification algorithms, especially Decision Trees, to analyze and predict the production and sales of freshwater fish in Batam. The method used is Decision Tree which involves collecting production data, preprocessing and analysis using RapidMiner. The developed model was tested with test data to assess its accuracy. The research results show that the classification model developed has 100% accuracy in predicting the production rate of perishable goods, such as freshwater fish, in Batam. This shows that the method used is effective in identifying and classifying fish types that have high potential for production, as well as providing accurate predictions for better business decision making.

KATA KUNCI:

Klasifikasi, Prediksi, Ikan Air Tawar, Produksi Ikan, Decison Tree, Batam

ABSTRAK

Manfaat penelitian ini mencakup aspek akademik, praktis, dan perilaku. Dari sudut pandang akademis, penelitian ini menambah literatur tentang aplikasi data mining di dunia nyata, khususnya dalam klasifikasi dan peramalan produksi minyak ikan. Secara praktis, penelitian ini menawarkan rekomendasi berbasis data kepada pemilik usaha perikanan di Batam untuk meningkatkan produktivitas dan efisiensi. Dari sudut pandang kebijakan, penelitian ini memberikan panduan bagi para pembuat kebijakan agar sumber daya air tawar perikanan sehari-hari dapat dimanfaatkan secara lebih tepat dan berkesinambungan. Penelitian di bidang ini memiliki keterbatasan tertentu. Awalnya penelitian hanya menggunakan data historis produksi ikan tuna di Batam. Lebih lanjut, fokus penelitian ini hanya sebatas pada klasifikasi dan prakiraan produksi ikan tawari saja, mengabaikan aspek lain seperti distribusi dan pemasaran. Metode ketiga yang digunakan dalam penelitian ini adalah metode data mining yang ada saat ini.

PENDAHULUAN

Batam, sebagai salah satu wilayah kepulauan yang strategis di Indonesia, memiliki potensi besar dalam industri perikanan, khususnya budidaya ikan air tawar. Dengan meningkatnya permintaan akan produk ikan air tawar baik dari pasar lokal maupun internasional, pelaku usaha perikanan di Batam menghadapi tantangan dalam meningkatkan efisiensi produksi dan penjualan [1]. Data yang besar dan kompleks yang dihasilkan dari aktivitas bisnis ini dapat memberikan wawasan yang berharga jika dianalisis dengan metode yang tepat. Dalam konteks ini, teknik data mining menjadi alat yang sangat penting untuk menggali informasi berharga yang dapat membantu pengambilan keputusan bisnis.

Perikanan air tawar, khususnya budidaya ikan, memiliki peran penting dalam mendukung perekonomian Indonesia, terutama di daerah Batam yang merupakan kawasan strategis dengan akses langsung ke pasar internasional. Budidaya ikan air tawar, seperti tilapia, memiliki potensi besar dalam memenuhi kebutuhan protein masyarakat dan mendukung ketahanan pangan nasional [1] Selain itu, penelitian juga menyoroti pentingnya analisis pendapatan dari budidaya ikan, seperti mujair, dalam mendukung ekonomi daerah dan meningkatkan pendapatan nelayan [2].

Penambangan data adalah proses yang berharga untuk menemukan pola yang berguna dari kumpulan data yang besar dan kompleks. Salah satu metode populer dalam penambangan data adalah penggunaan algoritma, seperti algoritma C4.5, yang mampu menangani data kontinu dan diskrit, memberikan hasil yang mudah dipahami oleh manusia [3] Algoritme ini cocok untuk menganalisis data penjualan terkait penjualan ikan air tawar di Batam, yang melibatkan variabel seperti spesies ikan, jumlah produksi, harga, dan waktu penjualan [2].

Klasifikasi ikan air tawar dan prediksi produksinya merupakan aspek penting dalam pengelolaan perikanan untuk meningkatkan efisiensi dan efektivitas[3]. Teknologi data mining memungkinkan analisis data historis produksi dan faktor-faktor lain yang mempengaruhi produksi dilakukan secara lebih akurat. Dalam konteks ini, penelitian tentang identifikasi dan klasifikasi spesies ikan air tawar menjadi krusial. Studi oleh [4] memberikan informasi tentang inventarisasi ikan di Sungai Kamikingi, yang dapat membantu dalam klasifikasi jenis-jenis ikan air tawar. Sementara itu, penelitian oleh [5] dan [6] yang menggunakan metode DNA barcoding dapat mendukung identifikasi spesies ikan secara akurat.

Untuk meningkatkan pengambilan keputusan di industri perikanan, penerapan teknik penambangan data seperti algoritma klasifikasi dan prediksi dapat memberikan rekomendasi yang lebih akurat bagi para pemangku kepentingan. Dengan memanfaatkan teknik ini, model dapat dikembangkan untuk membantu dalam berbagai aspek mulai dari perencanaan produksi hingga pemasaran tangkapan. Penambangan data melibatkan ekstraksi wawasan berharga dari kumpulan data besar, memungkinkan penemuan tren, pola, dan aturan [7]. Ini adalah proses yang bertujuan untuk mendapatkan informasi yang berguna dari sumber data yang luas [8]. Alat yang digunakan dalam penambangan data mencakup model statistik, algoritma matematika, dan metode pembelajaran mesin [9].

Pemanfaatan algoritma dalam analisis data penjualan ikan air tawar di Batam bertujuan untuk mengidentifikasi pola-pola penjualan yang signifikan dan faktor-faktor yang mempengaruhi peningkatan penjualan. Melalui analisis ini, pelaku usaha dapat memperoleh wawasan mendalam mengenai tren pasar, preferensi konsumen, serta faktor-faktor lain yang berdampak pada penjualan. Dengan demikian, strategi pemasaran yang lebih efektif dapat dirancang, yang pada akhirnya akan meningkatkan keuntungan dan keberlanjutan bisnis perikanan di Batam.

Penambangan data memainkan peran penting dalam meningkatkan efisiensi manajemen produksi dengan menganalisis pola permintaan, memungkinkan bisnis untuk menyelaraskan produksi ikan air tawar dengan kebutuhan pasar, sehingga mengurangi limbah dan biaya operasional. Melalui penerapan algoritma data mining, seperti algoritma Apriori, prediksi penjualan dapat dilakukan, membantu dalam peningkatan perencanaan bisnis [10]; [11]. Analisis prediktif ini memfasilitasi pengambilan keputusan yang lebih baik dalam hal perencanaan produksi dan strategi pemasaran, yang pada akhirnya mengarah pada manajemen logistik rantai pasokan dan gudang yang dioptimalkan [12] [13].

METODE PENELITIAN

Langkah pertama dalam proses analisis adalah menyiapkan kumpulan data yang akan digunakan untuk analisis. Setelah itu, lokasi atau direktori penyimpanan kumpulan data harus dipertimbangkan. Selanjutnya, setiap atribut dalam pengumpulan data harus disesuaikan jenisnya untuk memenuhi persyaratan analisis. Setelah kumpulan data ini diformat, kumpulan data tersebut siap untuk dianalisis secara lebih menyeluruh.

Selanjutnya dilakukan preprocessing yang disebut juga pembersihan data, seperti transformasi nilai yang hilang, analisis data yang tidak konsisten, atau normalisasi data. Setelah dilakukan preprocessing, ditentukan fungsi atau peran setiap atribut dalam kumpulan data, seperti atribut target dan prediktor.

Langkah selanjutnya adalah membagi kumpulan data menjadi data pelatihan dan pengujian. Meskipun pengujian data digunakan untuk menilai performa model, pelatihan data digunakan untuk menyempurnakan model. Algoritma Decision Tree digunakan untuk membuat model klasifikasi menggunakan data pelatihan yang telah diperiksa secara menyeluruh.

Setelah model dibuat, kemudian diuji menggunakan data uji atau data baru untuk melihat bagaimana kinerja model dalam kondisi dunia nyata. Performa model kemudian disesuaikan menggunakan metrik evaluasi yang relevan, seperti akurasi, tekanan, memori, dan lain-lain. Proses analisis ini diakhiri dengan langkah evaluasi kinerja yang menandakan bahwa proses analisis telah selesai dan dapat di lihat pada Gambar 1. Flowchart penelitian.



Gambar 1. Flowchart penelitian.

HASIL DAN PEMBAHASAN

Data sedang dilakukan saat ini. Langkah ini untuk mengumpulkan data yang akan digunakan dalam proses klasifikasi menggunakan alat Rapidminer. Rapid Miner menggunakan atribut pilihan operator untuk melakukan proses seleksi. Langkah pertama dalam proses ekstraksi data adalah membuat operator Excel yang digunakan untuk mengekstrak semua data yang disertakan dalam spreadsheet. Kumpulan data pertama memiliki 100 nilai dan 12 atribut. Lanjutkan dengan mengatur atribut operator pilih seperti yang ditunjukkan pada diagram di atas.

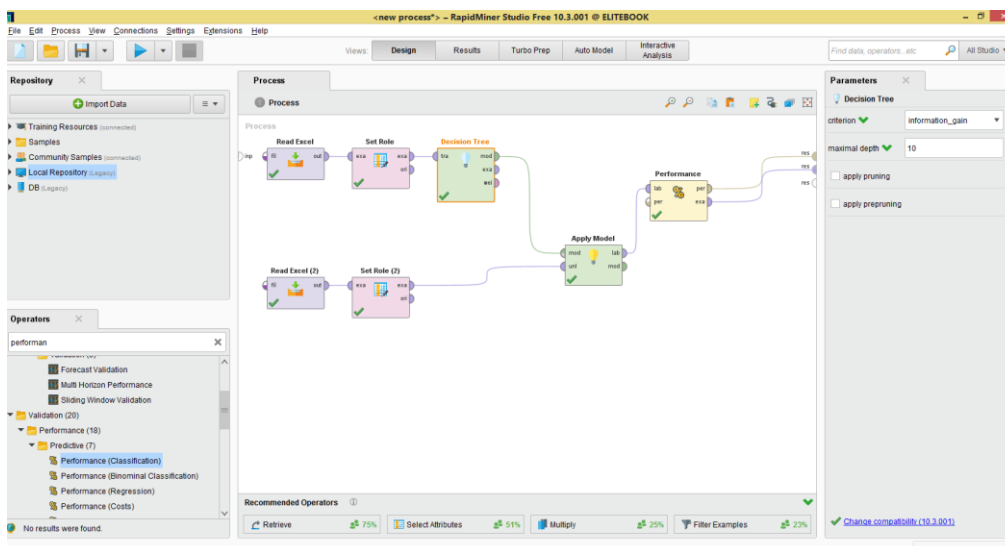
Operator membagi data untuk memisahkan data untuk pelatihan dan pengujian; 90% data diperoleh untuk pelatihan dan 10% untuk pengujian.

Model di atas merupakan model pohon keputusan. Dalam proses ini, operator pohon keputusan digunakan untuk menghasilkan fungsi probabilitas yang dapat diterapkan pada klasifikasi. Prosesnya dimulai dengan memilih operator penghapusan model yang berguna untuk menerapkan model tersebut.

Tabel 1 data set produksi perikanan dibatam

No	Jenis Ikan	Komoditi Ikan	Luas Kolam (m2)	Tinggi Kolam (m)	Jumlah Ikan	Jumlah Benih	Jumlah Pakan (kg)	Sumber Air	Produksi (kg)	Pembelian (Rp)	Harga Satuan (Rp/kg)
1	Lele	Lele Jumbo	110.02874	1.829547993	4297	7915	441.19684	Sumur	678.3878	2708327	29256.46223
2	Gurame	Lele Jumbo	163.5757357	2.615202554	4695	8797	381.34588	PDAM	701.2975	8919852	37788.04442
3	Nila	Lele Jumbo	83.36748087	1.257732231	4115	8755	103.71999	PDAM	1369.475	9443179	43896.83682
4	Gurame	Patin Lokal	60.90627065	1.081954434	3580	6560	443.02021	Sungai	1076.101	9116000	18528.4236
5	Patin	Gurame Super	143.590058	2.495617159	2150	7369	489.77515	Sungai	1217.018	4892298	49177.52125
6	Patin	Gurame Super	186.9519953	1.600920975	2328	8352	439.04977	PDAM	1752.249	9156231	21279.11807
7	Patin	Patin Lokal	194.0606537	1.204475789	4302	6078	264.56966	Sungai	544.1011	1893040	19537.88493
8	Gurame	Lele Jumbo	152.1714339	2.477951153	2589	5869	224.70664	PDAM	408.5987	7109114	30321.57423
9	Patin	Nila Merah	194.0159214	1.712951037	4577	7042	255.06792	Sungai	1913.597	6695442	38912.74462
10	Gurame	Patin Lokal	132.6210298	2.985004983	3164	6761	213.57071	PDAM	1726.938	2242759	43389.45407
11	Lele	Nila Merah	128.3399499	2.245660876	1025	9551	205.84152	Sungai	1514.983	6863463	28879.03109
12	Mas	Nila Merah	160.1242023	1.541216625	4540	8359	68.126871	Sungai	1700.291	3318977	20742.53448
13	Nila	Nila Merah	178.9645034	1.849973409	907	8530	115.01821	PDAM	1844.373	2436397	18837.28571
14	Mas	Patin Lokal	129.8506377	1.531577586	4672	1208	86.686142	Sungai	1969.722	8912648	39051.47455
15	Nila	Mas Koki	76.69692904	1.174035548	3659	2537	369.03592	PDAM	1198.956	3543559	48590.56251
16	Gurame	Patin Lokal	158.3910494	2.663165342	3482	2555	496.14743	Mata Air	513.2178	8766161	48123.30004
17	Gurame	Lele Jumbo	165.0277871	2.507445109	1449	3663	133.97186	Sungai	1676.006	9879199	32912.71778
18	Patin	Patin Lokal	193.3699044	1.445497089	1537	2260	129.16697	Mata Air	1468.355	6257737	31311.43219
19	Patin	Nila Merah	113.8667621	1.790945713	1149	4210	220.61843	Sungai	1789.037	7019466	23518.14246
20	Lele	Patin Lokal	61.25899647	2.95740205	2154	3967	109.80998	PDAM	740.8485	2309230	39296.61385
21	Nila	Nila Merah	92.67627225	1.374909034	4744	4935	310.50078	Sungai	1888.473	4702640	46567.3045
22	Patin	Gurame Super	79.84331238	2.246876429	4768	3437	54.385799	Sumur	1578.437	9906530	24948.0883
23	Gurame	Nila Merah	147.3016652	2.933493966	1807	1566	379.42896	Sungai	1110.99	9596728	42711.57071
24	Gurame	Patin Lokal	172.3871925	2.23878713	1095	8537	461.02138	PDAM	866.5872	2371563	16475.82384
25	Gurame	Mas Koki	172.9499851	2.891324956	1967	5404	68.500756	Mata Air	1488.537	1026914	35716.93754

RAPIDMINER

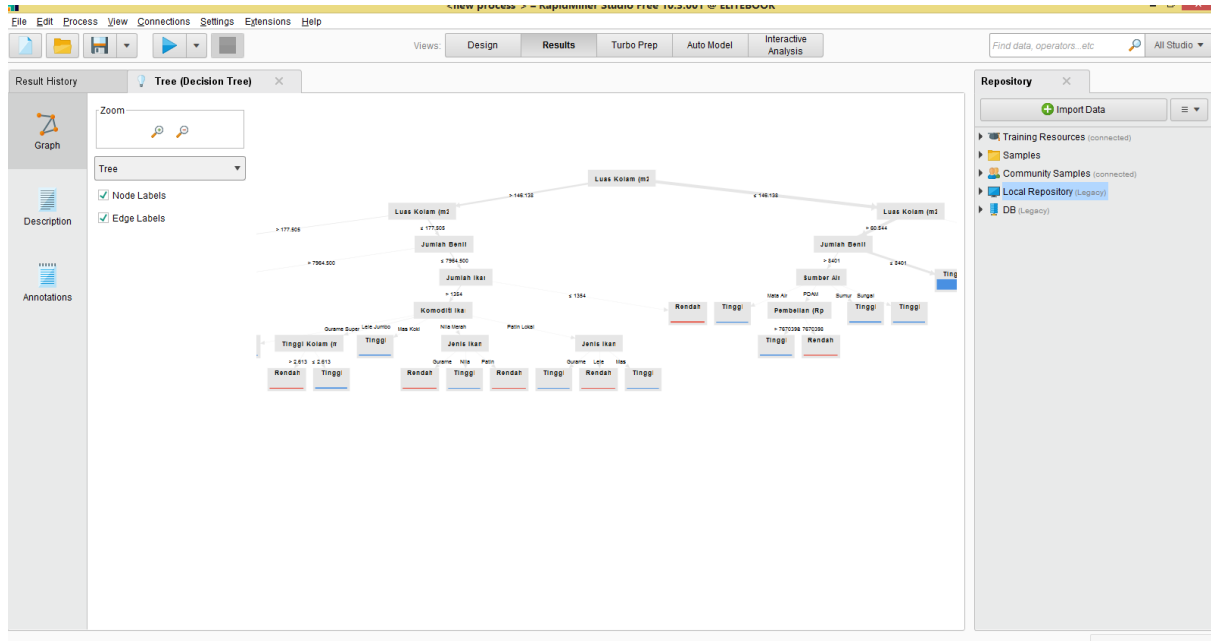


Gambar 2 Operator pada Rapidminer

Pada tahap ini dilakukan seleksi data. Langkah ini untuk mengumpulkan data yang akan digunakan dalam proses klasifikasi menggunakan alat Rapidminer. Atribut pilihan operator digunakan di Rapid Miner untuk melakukan proses seleksi. Proses pemilihan data dimulai dengan membuat operator Excel yang dimaksudkan untuk mengekstrak semua data yang ada di spreadsheet, dimana data pertama memiliki 100

nilai dan 12 atribut. Lanjutkan dengan mengatur atribut operator pilih seperti yang ditunjukkan pada gambar diatas.

DECISION TREE



Gambar 3 Hasil Decision Tree

Model diatas merupakan model pohon keputusan. Dalam proses ini operator pohon keputusan digunakan untuk menghasilkan fungsi probabilitas yang dapat diterapkan pada klasifikasi. Prosesnya dimulai dengan memilih operator penghapusan model yang berguna untuk menerapkan model tersebut.

Tree

```

Jumlah Pakan (kg) > 489.992: Rendah {Tinggi=1, Rendah=2}
Jumlah Pakan (kg) ≤ 489.992
| Tinggi Kolam (m) > 1.034
| | Jumlah Pakan (kg) > 62.997
| | | Tinggi Kolam (m) > 2.707
| | | | Pembelian (Rp) > 6464982: Rendah {Tinggi=1, Rendah=4}
| | | | Pembelian (Rp) ≤ 6464982: Tinggi {Tinggi=9, Rendah=0}
| | | Tinggi Kolam (m) ≤ 2.707
| | | | Jumlah Ikan > 1658.500: Tinggi {Tinggi=51, Rendah=1}
| | | | Jumlah Ikan ≤ 1658.500
| | | | | Tinggi Kolam (m) > 2.022
| | | | | | Tinggi Kolam (m) > 2.054: Tinggi {Tinggi=10, Rendah=2}
| | | | | | Tinggi Kolam (m) ≤ 2.054: Rendah {Tinggi=0, Rendah=2}
| | | | | Tinggi Kolam (m) ≤ 2.022: Tinggi {Tinggi=11, Rendah=0}
| | Jumlah Pakan (kg) ≤ 62.997: Tinggi {Tinggi=1, Rendah=1}
| Tinggi Kolam (m) ≤ 1.034
| | Komoditi Ikan = Gurame Super: Tinggi {Tinggi=2, Rendah=0}
| | Komoditi Ikan = Patin Lokal: Rendah {Tinggi=0, Rendah=2}

```

Gambar 4. Tree yang di hasilkan

Performance

accuracy: 100.00%

	true Tinggi	true Rendah	class precision
pred. Tinggi	65	0	100.00%
pred. Rendah	0	35	100.00%
class recall	100.00%	100.00%	

Gambar 5. Performance

Keakuratan data yang diekstrak dari set pelatihan adalah 100,00%. Grafik di atas menunjukkan tinggi yang diprediksi dengan tinggi asli di atas 65 dan rendah sebenarnya 0. Dapat dilihat bahwa pengembalian yang diprediksi lebih dari nol ketika pengembalian sebenarnya adalah 35. Baik tinggi sebenarnya maupun rendah sebenarnya memiliki hasil yang sama dalam hal penarikan data.

KESIMPULAN DAN SARAN

Berdasarkan evaluasi model dan uji coba dapat disimpulkan bahwa pengklasifikasian tingkat produksi barang mudah rusak dengan algoritma dibatam cukup efektif .

Akurasi algoritma dalam memprediksi ambang produksi perikanan di Kabupaten Sukabumi mencapai 100,00%.

Hal ini menunjukkan bahwa model yang dikembangkan mempunyai kemampuan dalam memberikan prediksi yang sangat akurat mengenai laju produksi barang yang mudah rusak.

Dengan memanfaatkan algoritma, penelitian ini diharapkan dapat mengembangkan model klasifikasi yang efektif untuk mengidentifikasi jenis ikan tawari yang memiliki potensi produksi tinggi di Batam. Model ini akan membantu pengguna perikanan memberikan respons yang lebih akurat dan efisien dan juga akan meminimalkan kebutuhan akan perikanan yang sedang berlangsung.

DAFTAR PUSTAKA

- [1] A. A. Dermawan, "Analisis Faktor-faktor Pengaruh Keputusan Pembelian Produk Kosmetik Skin Care melalui Offline dan Online," 2020. [Online]. Available: <https://repositori.usu.ac.id/handle/123456789/28331>
- [2] A. Agung Dermawan, H. Nasution, and M. Haikal Sitepu, "The impact of branding on purchasing decision-making in mall shopping and online shopping," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 801, no. 1, 2020, doi: 10.1088/1757-899X/801/1/012146.
- [3] M. Muqimuddin, A. A. Darmawan, and B. N. Abdallah, "Prioritas Penyelesaian Akar Masalah Kualitas Palm Kernel Oil Dengan Memperhatikan Uncertain Information," *J. Optimasi Tek. Ind.*, vol. 4, no. 2, p. 51, 2022, doi: 10.30998/joti.v4i2.13631.
- [4] Kaunang, S., "Analysis of income from mujair fish (*Oreochromis massambicus*) cultivation in tarpon ponds (biofloc) (a case study of biofloc business in Matungkas village, Minut regency)," *Devotion Journal of Research and Community Service*, vol. 4, no. 12, pp. 2253-2262, 2023. doi: 10.59188/devotion.v4i12.625.
- [5] Liufeto, F., "Strategies in developing freshwater fish farming to prevent stunting," *Interdisciplinary Social Studies*, vol. 2, no. 4, pp. 1905-1912, 2023. doi: 10.55324/iss.v2i4.416.
- [7] Fadhila, F. and Hasugian, P., "Application of C4.5 algorithm to prediction sales at PT. Sumber Sayur Segar," *Journal of Intelligent Decision Support System (IDSS)*, vol. 5, no. 1, pp. 10-19, 2022. doi: 10.35335/idss.v5i1.45.
- [8] Ahmed, S., Datta, S., and Zhilik, A., "Molecular diversity of freshwater fishes of Bangladesh assessed by DNA barcoding," *Bangladesh Journal of Zoology*, vol. 48, no. 1, pp. 1-19, 2020. doi: 10.3329/bjz.v48i1.47872.
- [9] Kalindiro, R., "Ichthyological inventory of the Kamikingi river: Identification and classification of predominant species," *Environmental and Earth Sciences Research Journal*, vol. 10, no. 3, pp. 93-99, 2023. doi: 10.18280/eesrj.100302.
- [10] Tsoupas, A., Papavasileiou, S., Minoudi, S., Gkagkavouzis, K., Petriki, O., Bobori, D., et al., "DNA barcoding identification of Greek freshwater fishes," *PLOS ONE*, vol. 17, no. 1, pp. e0263118, 2022. doi: 10.1371/journal.pone.0263118.

- [11] Dabab, M., Freiling, M., Rahman, N., and Sagalowicz, D., "A decision model for data mining techniques," *Proceedings of PICMET '18 Conference*, pp. 1-5, 2018. doi: 10.23919/picmet.2018.8481953.
- [12] R, A. and SashiKumar, D., "A survey of data mining techniques for internet," *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 7, no. 1, pp. 22-35, 2017. doi: 10.23956/ijarcsse/v7i1/0125.
- [13] Rahim, S., Rahman, M., and Chowdhury, A., "Mining industrial engineered data of apparel industry: A proposed methodology," *International Journal of Computer Applications*, vol. 161, no. 7, pp. 1-7, 2017. doi: 10.5120/ijca2017913262.