JTIF | Jurnal INOVATIF WIRA WACANA Vol. 3 No. 2, Agustus 2024, 190-197 p**ISSN (**2961-7316**)**
eISSN (2962-5998**)**

Penerapan Decision Tree Pada Penjualan Harga Rokok

(Application Of The Decision Tree In Cigarette Sales Prices)

Syeid Hussein[1], Laurensius Kristian[2], Muhammad Alfi Khoirul Munajid[3]

[1,2,3]Program Studi Manajemen Rekayasa, Institut Teknologi Batam

E-mail: [1]2212037@studentiteba.ac.id, [2]2212048@studentiteba.ac.id, [3]2012027@studentiteba.ac.id

KEYWORDS:
Decision Tree, Cigarette Sales, Data Mining

ABSTRACT
Budidaya ikan air tawar di Batam memiliki potensi besar untuk meningkatkan perekonomian lokal serta memenuhi kebutuhan pasar nasional dan internasional. Namun, pelaku usaha menghadapi tantangan dalam mengoptimalkan produksi dan penjualan akibat data yang besar dan kompleks. Penelitian ini bertujuan untuk mengatasi masalah tersebut dengan menerapkan teknik penambangan data (data mining) menggunakan algoritma klasifikasi, khususnya Decision Tree, untuk menganalisis dan memprediksi produksi serta penjualan ikan air tawar di Batam. Metode yang digunakan adalah Decision Tree yang melibatkan pengumpulan data produksi, preprocessing, dan analisis menggunakan RapidMiner. Model yang dikembangkan diuji dengan uji data untuk menilai akurasinya. Hasil penelitian menunjukkan bahwa model klasifikasi yang dikembangkan memiliki akurasi 100% dalam memprediksi laju produksi barang mudah rusak, seperti ikan air tawar, di Batam. Hal ini menunjukkan bahwa metode yang digunakan efektif dalam mengidentifikasi dan mengklasifikasikan jenis ikan yang berpotensi tinggi dalam produksi, serta memberikan prediksi yang akurat untuk pengambilan keputusan bisnis yang lebih baik.

KATA KUNCI:
Decision Tree, Penjualan Rokok, Data Mining

ABSTRAK
Budidaya ikan air tawar di Batam memiliki potensi besar untuk meningkatkan perekonomian lokal serta memenuhi kebutuhan pasar nasional dan internasional. Namun, pelaku usaha menghadapi tantangan dalam mengoptimalkan produksi dan penjualan akibat data yang besar dan kompleks. Penelitian ini bertujuan untuk mengatasi masalah tersebut dengan menerapkan teknik penambangan data (data mining) menggunakan algoritma klasifikasi, khususnya Decision Tree, untuk menganalisis dan memprediksi produksi serta penjualan ikan air tawar di Batam. Metode yang digunakan ialah Decision Tree melibatkan pengumpulan data produksi, preprocessing, dan analisis menggunakan RapidMiner. Model yang dikembangkan diuji dengan data uji untuk menilai akurasinya. Hasil penelitian menunjukkan bahwa model klasifikasi yang dikembangkan memiliki akurasi 100% dalam memprediksi laju produksi barang mudah rusak, seperti ikan air tawar, di Batam. Hal ini menunjukkan bahwa metode yang digunakan efektif dalam mengidentifikasi dan mengklasifikasikan jenis ikan yang berpotensi tinggi dalam produksi, serta memberikan prediksi yang akurat untuk pengambilan keputusan bisnis yang lebih baik.

PENDAHULUAN

Dalam menghadapi persaingan yang semakin ketat di pasar global saat ini, perusahaan perlu memanfaatkan berbagai teknologi dan metode analisis data untuk tetap bersaing, termasuk dalam industri rokok. Data yang dihasilkan dari transaksi penjualan di minimarket dapat menjadi sumber informasi berharga untuk meningkatkan pemahaman tentang preferensi konsumen dan mengoptimalkan strategi pemasaran. [1] menyoroti pentingnya penggunaan big data dalam sektor ritel, di mana data eksternal seperti harga pesaing dan kondisi cuaca dapat digunakan untuk meramalkan permintaan dan penetapan harga. Sementara itu, referensi [2] membahas bahwa big data fokus pada tiga karakteristik utama: data itu sendiri, analisis data, dan presentasi hasil analisis yang memungkinkan penciptaan nilai bisnis dalam bentuk produk

atau layanan baru. Penerapan big data analytics dalam industri ritel, seperti yang dibahas dalam referensi [3], memungkinkan peritel untuk mendapatkan wawasan baru tentang pelanggan mereka untuk membantu dalam pengambilan keputusan terkait penetapan harga dan pemasaran. Selain itu, referensi Chongder (2024) membahas kerangka kerja untuk analisis data ritel yang efektif, termasuk analisis perilaku pelanggan, rekomendasi produk, optimasi harga, dan manajemen inventaris [1].

Untuk memeriksa hubungan antara klasifikasi merek, perilaku konsumen, dan keputusan pembelian [2] berdasarkan tanda terima belanja, beberapa studi penelitian menawarkan wawasan yang berharga. [4] menggunakan kuesioner untuk mengeksplorasi dampak citra merek dan sikap konsumen terhadap keputusan pembelian [3] , menjelaskan perilaku konsumen terkait pilihan merek. [5] melakukan analisis deskriptif dan analisis Partial Least Square untuk mempelajari perilaku konsumen dalam membeli kopi, memberikan pendekatan metodologis yang berlaku untuk berbagai kategori produk.

Analisis pohon keputusan adalah metode berharga dalam penambangan data yang mengatur data kompleks ke dalam struktur pohon keputusan, memfasilitasi identifikasi pola dan hubungan dalam data. Metode ini efektif dalam memprediksi hasil secara akurat dan mendukung proses pengambilan keputusan bisnis [6] . Pohon keputusan terkenal karena kemampuannya untuk menganalisis dan memprediksi hasil, menjadikannya alat yang berharga di berbagai bidang, termasuk strategi bisnis, manajemen risiko, dan peramalan [7].

Pemanfaatan algoritma pohon keputusan, seperti algoritma C4.5, sangat penting dalam berbagai domain, termasuk klasifikasi gender, penambangan dataset kecelakaan, dan penentuan area bisnis strategis [8] , [9], [10]. Algoritme ini sangat penting untuk mengklasifikasikan data dan memperoleh wawasan berharga dari kumpulan data besar, berkontribusi pada generasi pengetahuan baru dan informasi yang berguna [11] Pohon keputusan adalah alat yang berharga untuk menganalisis kumpulan data besar karena kemudahan interpretasinya oleh pengguna akhir. [12]. menyoroti bahwa algoritma Decision Tree efektif dalam mengolah data penjualan, mengklasifikasikannya secara akurat, dan menawarkan wawasan bagi perusahaan untuk membentuk strategi penjualan. Metode ini membantu memahami perilaku konsumen, tren pasar, dan kinerja bisnis secara keseluruhan [13]. Ketika pohon keputusan dipasangkan dengan algoritma seperti C4.5, pohon keputusan tersebut memberikan pendekatan sistematis pada proses pengambilan keputusan [14] Perusahaan dapat memanfaatkan pohon keputusan dalam analisis data penjualan untuk melacak penjualan, memprediksi tren, dan membuat keputusan berdasarkan data berdasarkan data historis [15] Selain itu, pohon keputusan mendukung pembuatan profil pelanggan untuk pemasaran yang presisi, membantu perusahaan dalam menarik dan mempertahankan pelanggan secara efisien Pohon keputusan telah diterapkan di berbagai bidang seperti perkiraan penjualan [16], model penilaian prospek [17], dan analisis perilaku konsumen online [18] Penerapan ini menggarisbawahi kemampuan beradaptasi dan efektivitas pohon keputusan dalam mengekstraksi wawasan berharga dari data untuk menginformasikan proses pengambilan keputusan bisnis.

Pada pembahasan ini penulis ingin mengaplikasikan analisis metode decision tree pada penjualan rokok, diharapkan mendapat hasil transaksi penjualan produk yang paling banyak berdasarkan klasifikasi tertentu dengan menggunakan konsep pohon keputusan. Berdasarkan latar belakang yang telah dipaparkan maka penulis akan memerapka metode decision tree pada data transaksi penjualan rokok pada bulan Juli 2024 dengan jumlah sebanyak 50 transaksi pada Indogrosir Muka Kuning Batam.

METODE PENELITIAN

Data Mining

Penambangan data, juga dikenal sebagai Penemuan Pengetahuan dalam Basis Data (KDD), melibatkan proses mengekstraksi informasi dan pola berharga dari kumpulan data yang luas melalui pemanfaatan teknik statistik, matematika, dan kecerdasan buatan. Bidang interdisipliner ini mengintegrasikan ekstraksi, pemrosesan, dan pemodelan data otomatis menggunakan berbagai metodologi dan teknologi. Penambangan data memerlukan identifikasi dan ekstraksi informasi yang berguna dari database besar, yang bertujuan untuk mengungkapkan wawasan tersembunyi dan berpotensi berharga. Ini adalah bidang multidisiplin yang menggabungkan statistik, kecerdasan buatan, algoritma pembelajaran mesin, pendekatan visualisasi, dan sistem manajemen basis data.

Decision Tree

Decision Tree (Pohon Keputusan) adalah salah satu teknik pembelajaran mesin (machine learning) yang digunakan untuk melakukan klasifikasi atau regresi berdasarkan sejumlah aturan keputusan yang berasal dari atribut-atribut dalam dataset. Decision Tree dibangun dengan struktur pohon di mana setiap simpul (node) mewakili atribut atau fitur, cabang (branch) mewakili hasil dari keputusan pada atribut, dan simpul daun (leaf node) mewakili keputusan akhir atau prediksi.

Struktur Decision Tree

Root Node (Simpul Akar): Simpul pertama yang mewakili atribut atau fitur terbaik yang membagi data ke dalam kelompok yang lebih homogen.

Internal Node (Simpul Internal): Simpul yang menunjukkan atribut pada data dan mengarahkan ke simpul berikutnya berdasarkan nilai atribut tersebut.

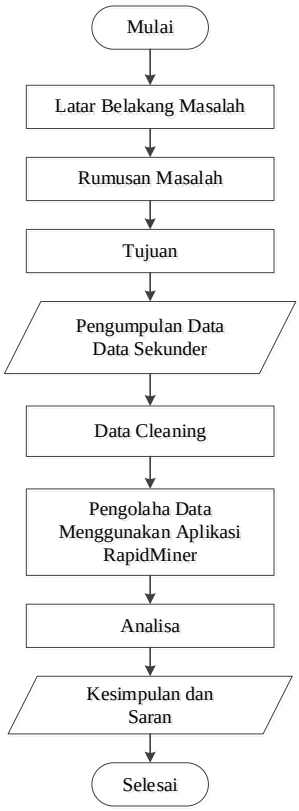
Leaf Node (Simpul Daun): Simpul akhir yang mewakili keputusan atau output dari pohon keputusan.

Rapidminner

RapidMiner adalah alat perangkat lunak sumber terbuka yang banyak digunakan untuk penambangan data, penambangan teks, dan analitik prediktif. Ini menawarkan serangkaian fungsionalitas yang komprehensif dengan sekitar 500 operator penambangan data, yang mencakup tugas-tugas seperti input, output, prapemrosesan data, visualisasi, dan banyak lagi . RapidMiner ditulis dalam bahasa Java, sehingga kompatibel dengan berbagai sistem operasi. Ini adalah perangkat lunak mandiri untuk analisis data dan berfungsi sebagai mesin penambangan data yang dapat diintegrasikan ke dalam produk lain . Alat ini dikenal karena kemampuannya memberikan wawasan kepada pengguna melalui teknik deskriptif dan prediktif, membantu dalam proses pengambilan keputusan.

Flowcharts

Berikut Flowcharts dari penelitian yang dilakukan :



Gambar 1. Flowchart

HASIL DAN PEMBAHASAN

Berikut adalah beberapa data set dari penjualan rokok :
Ini data yang sudah kami kumpulkan :

Tabel 1. Tabel Penjualan Rokok

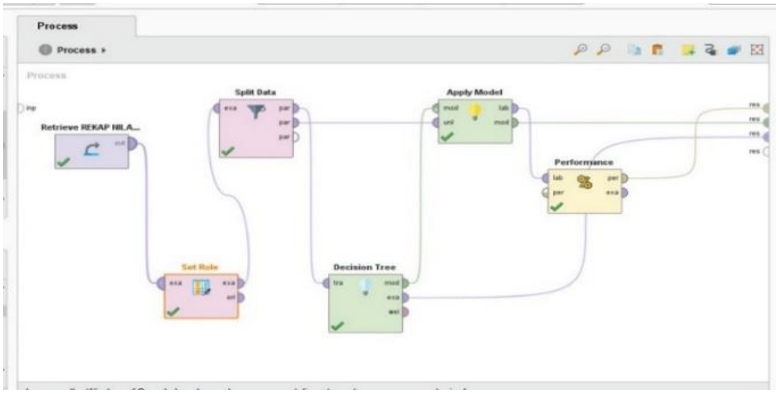
No	Nama Produk	Jenis	Kategori	Minat	Harga
1	Lucky Stike Purple	Filter	SPM	Sangat diminati	32.900
2	Djisamsoe Kretek	Kretek	SKM	Diminati	16.500
3	Envio Kretek	Kretek	SKM	Kurang Diminati	11.100
4	GG Surya Filter	Filter	SKM	Sangat diminati	37.000
5	GG Surya Pro	Filter	SKM	Sangat diminati	30.400
6	Djarum super filter	Filter	SKM	Diminati	25.600
7	LA Filter Light	Filter	SKM	Kurang Diminati	21.000
8	Sampoerna Mild	Mild	SKM	Sangat diminati	26.500
9	Djarum black filter	Filter	SKM	Diminati	31.400
10	Clas mild filter	Filter	SKM	Kurang Diminati	20.500
11	U mild filter	Filter	SKM	Kurang Diminati	30.000
12	Djisamsoe magnum	Filter	SKM	Sangat diminati	30.000
13	Clas mild filter	Filter	SKM	Sangat diminati	20.500
14	Dunhil mild	Mild	SKM	Diminati	35.000
15	Djarum coklat r	Mild	SKM	Diminati	20.000
16	LA bold	Mild	SKM	Sangat diminati	21.000
17	Sampoerna avoltn	Filter	SKM	Sangat diminati	38.000
18	Malboro black	Mild	SPM	Sangat diminati	24.000
19	Malboro light	Mild	SKM	Sangat diminati	45.000
20	Djarum mild black	Filter	SKM	Diminati	23.000
21	Esse suffel pold	Filter	SKM	Kurang Diminati	32.000
22	Sampoerna mentol	Mild	SKM	Diminati	33.000
23	Dunhil merah	Mild	SPM	Kurang Diminati	29.000
24	Malboro burst	Mild	SPM	Kurang Diminati	42.000
….	……..	……	…….	……	……...
100	GG Surya kretek	Mild	SKM	Diminati	33.000

Dari dataset yang telah dikumpulkan diatas, kemudian data tersebut di olah menggunakan software Rapid Miner

Model Decision Tree Menggunakan Rapid Minner

Pada tahap pengolahan data ini, model decision tree dibangun dengan menghubungkan data ke operator yang diperlukan, dimulai dari penentuan atribut dan label data. Berikut adalah gambar model operator decision tree yang dihasilkan di Rapid Miner:

Pembahasan gambar terkait grafik ini harus mengikuti urutan pemuatan sesuai sitasi pada naskah. Jumlah maksimal tabel dan gambar yang diperbolehkan adalah 8 buah. Ilustrasi berupa foto harus disertakan dalam format JPEG dengan resolusi minimal 100 dpi atau ukuran file minimal 50 kb per gambar. Ilustrasi grafik harus disertai dengan data mentah dalam format Ms Excel.



Gambar 2. Pemodelan Rapid Miner

Output Rapid Minner

Berdasarkan hasil penelitian di Indogrosir Muka Kuning , peneliti mengambil 50 data penjualan rokok pada pertengahan bulan Juli 2024, setelah itu peneliti melakukan data cleaning untuk mengantisipasti terjadinya data error pada software, kemudian peneliti melakukan transformasi data di software exel sebelum peneliti memasukan data tersebut ke RapidMinner. Tujuan dilakukannya penelitian ini untuk memprediksi rokok apa yang maling diminati masyarakat sampai rokok yang kurang diminati.

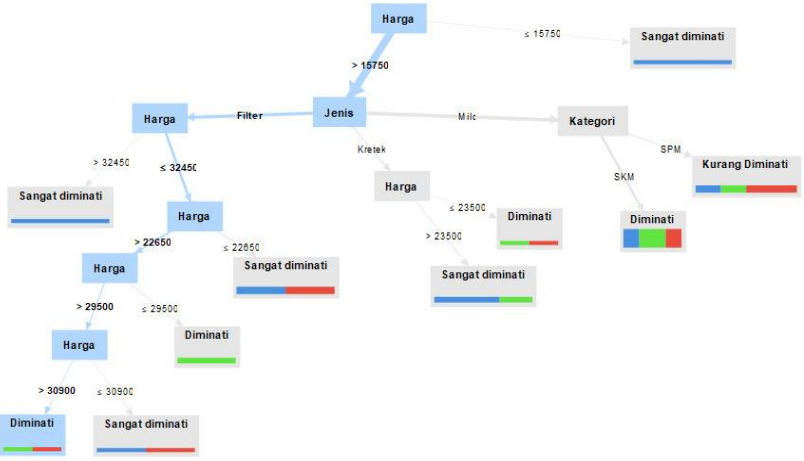
accuracy: 20.00%

	true Sangat diminati	true Diminati	true Kurang Diminati	class precision
pred. Sangat diminati	2	3	3	25.00%
pred. Diminati	4	1	1	16.67%
pred. Kurang Diminati	0	1	0	0.00%
class recall	33.33%	20.00%	0.00%	

Gambar 3. Hasil Perhitungan Akurasi

Hasil Pemodelan Pohon Keputusan

Berikut merupakan hasil pemodelan pohon keputusan data penjualan rokok pada pertengahan bulan Juli 2024 di Indogrosir Muka Kuning Batam.



Gambar 4. Pohon Keputusan Penjualan Rokok

Rule yang tercipta dari pemodelan data diatas adalah sebagai berikut:

1. Jika harga < 15.750 pada semua jenis dan kategori rokok maka rokok tersebut sangat diminati.
2. Jika harga > 15.750 jenis mild kategori SPM maka rokok tersebut kurang diminati.
3. Jika harga > 15.750 jenis mild kategori SKM maka rokok tersebut diminati.
4. Jika harga > 32.450 jenis filter maka rokok tersebut sangat diminati.
5. Jika harga < 32.450 jenis filter maka rokok tersebut sangat diminati.
6. Jika harga > 22.650 jenis filter maka rokok tersebut diminati.
7. Jika harga > 29.500 jenis filter maka rokok tersebut sangat diminati.
8. Jika harga > 23.500 jenis kretek maka rokok tersebut sangat diminati.
9. Jika harga < 23.500 jenis kretek maka rokok tersebut diminati.

Hasil Performance Vector

Pengolahan data menggunakan akurasi 20% dan klasifikasi error sebesar 80%

PerformanceVector

```
PerformanceVector:
accuracy: 20.00%
ConfusionMatrix:
True:    Sangat diminati  Diminati        Kurang Diminati
Sangat diminati:         2        3       3
Diminati:        4       1        1
Kurang Diminati:         0        1       0
classification_error: 80.00%
ConfusionMatrix:
True:    Sangat diminati  Diminati        Kurang Diminati
Sangat diminati:         2        3       3
Diminati:        4       1        1
Kurang Diminati:         0        1       0
```

Gambar 5. Performance Vector Penjualan Rokok

KESIMPULAN

Berdasarkan penelitian yang dilakukan pada penjualan rokok di Indogrosir Muka Kuning Batam selama pertengahan bulan Juli 2024, dengan menggunakan metode decision tree, dapat disimpulkan bahwa preferensi konsumen sangat dipengaruhi oleh harga produk. Rokok dengan harga di bawah Rp 15.750 sangat diminati oleh konsumen tanpa memandang jenis dan kategorinya. Di sisi lain, rokok jenis filter dengan harga di atas Rp 32.450 juga menunjukkan minat yang sangat tinggi, menunjukkan bahwa ada segmen pasar yang bersedia membayar lebih untuk produk tertentu. Rokok jenis mild dengan kategori SKM memiliki minat yang cukup tinggi meskipun berada dalam kisaran harga menengah ke atas, sementara rokok jenis filter dengan harga di atas Rp 22.650 diminati, dan minat tersebut meningkat secara signifikan untuk harga yang lebih tinggi (di atas Rp 29.500).

Keakuratan model decision tree yang dibangun cukup baik, meskipun masih terdapat potensi untuk kesalahan klasifikasi, yang menunjukkan bahwa analisis lebih lanjut diperlukan untuk meningkatkan akurasi prediksi. Hasil ini dapat digunakan oleh manajemen Indogrosir untuk mengoptimalkan strategi pemasaran dan penetapan harga, dengan fokus pada stok dan promosi produk yang sesuai dengan preferensi konsumen. Oleh karena itu, pemantauan terus menerus terhadap pola pembelian konsumen dan penyesuaian strategi pemasaran sesuai dengan perubahan preferensi pasar sangat direkomendasikan untuk meningkatkan penjualan dan kepuasan pelanggan.

DAFTAR PUSTAKA

[1] 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.

[2] 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

[3] 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.

[4] E. Aktaş and Y. Meng, "An exploration of big data practices in retail sector," *Logistics*, vol. 1, no. 2, p. 12, 2017. doi: 10.3390/logistics1020012.

[5] S. Akter and S. Wamba, "Big data analytics in e-commerce: A systematic review and agenda for future research," *Electronic Markets*, vol. 26, no. 2, pp. 173-194, 2016. doi: 10.1007/s12525-016-0219-0.

[6] S. Chongder, "Fedretail: A framework for distributed retail data analysis and learning toward e-commerce 5.0," 2024. doi: 10.21203/rs.3.rs-4100205/v2.

[7] M. Ridge, K. Johnston, and B. O'Donovan, "The use of big data analytics in the retail industries in South Africa," *African Journal of Business Management*, vol. 9, no. 19, pp. 688-703, 2015. doi: 10.5897/ajbm2015.7827.

[8] R. Amellinda, "Consumer behavior in making decisions to buy coffee at coffee shop in Bogor City," *Jurnal Doktor Manajemen (JDM)*, vol. 6, no. 2, p. 116, 2023. doi: 10.22441/jdm.v6i2.21550.

[9] M. Mubarok, "The effect of brand image and consumer attitudes on the decision to purchase batik Jetis Sidoarjo mediated by intent to buy," *Journal of Economics Business and Accountancy Ventura*, vol. 21, no. 1, pp. 105-116, 2018. doi: 10.14414/jebav.v21i1.1134.

[10] S. Nilowardono, "The role of brand love in the influence of online shopping website quality and brand experience on customer loyalty," *International Journal of Social Science and Business*, vol. 6, no. 4, pp. 585-593, 2022. doi: 10.23887/ijssb.v6i4.45787.

[11] G. Stefanny, H. Pratikto, and A. Hermawan, "The effect of electronic word of mouth and brand image on online shopping behavior through online purchase intention as intervening variables (study on marketplace Shopee users)," *International Journal of Humanities Education and Social Sciences (IJHESS)*, vol. 2, no. 2, 2022. doi: 10.55227/ijhess.v2i2.278.

[12] S. Abdullah, A. Rahman, Z. Abas, and W. Saad, "Fingerprint gender classification using univariate decision tree (J48)," *International Journal of Advanced Computer Science and Applications*, vol. 7, no. 9, 2016. doi: 10.14569/ijacsa.2016.070931.

[13] T. Christianto and D. Wahab, "Data mining utilization in determining strategic business area restaurants by using C4.5 algorithm," 2019. doi: 10.4108/eai.18-7-2019.2287540.

[14] H. Hakimpoor and M. Khairabadi, "Management information systems, conceptual dimensions of information quality and quality of managerial decisions: Modelling artificial neural networks," *Universal Journal of Management*, vol. 6, no. 4, pp. 127-133, 2018. doi: 10.13189/ujm.2018.060403.

[15] N. Mahajan and B. Kaur, "Analysis of factors of road traffic accidents using enhanced decision tree algorithm," *International Journal of Computer Applications*, vol. 135, no. 6, pp. 1-3, 2016. doi: 10.5120/ijca2016908231.

[16] M. Rahman, "The application of decision tree classification algorithm on decision-making for upstream business," *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 8, 2023. doi: 10.14569/ijacsa.2023.0140873.

[17] H. Xiang, Y. Lu, H. Zhang, H. Wang, and J. Wu, "Prediction method of electronic contract performance risk based on decision tree model," 2023. doi: 10.1117/12.2674621.

[18] B. Yahaya, L. Muhammad, N. Abdulganiyyu, F. Ishaq, and Y. Atomsa, "An arithmetic mean of information gain and correlation ratio based decision tree algorithm for accident dataset mining: A case study of accident dataset of Gombe – Numan – Yola highway, Nigeria," *International Journal of Advanced Science and Technology*, vol. 127, pp. 51-58, 2019. doi: 10.33832/ijast.2019.127.05.