

PENERAPAN ALGORITMA *DECISION TREE* DALAM PENJUALAN *HANDPHONE*

Faisal¹, Harry Dhika², Helmi Veris³

^{1,2,3}Program Studi Teknik Informatika, Fakultas Teknik dan Ilmu Komputer,
Universitas Indraprasta PGRI

Jalan Raya Tengah No 80, Kelurahan Gedong, Pasar Rebo, Jakarta Timur
faisalnin7@gmail.com¹, dhikatr@yahoo.com², helmiveris@gmail.com³

Abstrak

Penelitian ini melakukan penerapan algoritma decision tree. perusahaan kurang menganalisa penjualan barang dan stok persediaan barang didalam gudang. jadi jenis penelitian ini yakni memprediksi dengan konsep data mining dengan sample data penjualan bulan lalu. Tujuan penelitian ini untuk mengetahui produk apa yang paling diminati konsumen dan produk apa yang harus direstock, agar persediaan barang terkontrol. Metode penelitian yang akan digunakan untuk malakukan penerapan algoritma decision tree ini adalah menggunakan metode kualitatif, dan metode pengumpulan data-data dengan cara wawancara, studi kepustakaan, dan observasi. hasil penelitian terdapat tingkat akurasi sebesar 88,89% dengan nilai presisi pada prediksi restok 87,50% dan nilai presisi pada prediksi tidak restok 100%. lalu untuk class recall restok 100% dan untuk class recall tidak restok 50%.

Kata Kunci :Algoritma Decision Tree, Penjualan Handphone, dan Data Mining.

Abstract

This study applies the decision tree algorithm. the company does not analyze the sale of goods and inventory stock in the warehouse. so this type of research is predicting with the concept of data mining with a sample of last month's sales data. The purpose of this study is to find out what products are most in demand by consumers and what products must be restocked, so that inventory is controlled. The research method that will be used to implement the decision tree algorithm is using qualitative methods, and methods of collecting data by means of interviews, literature study, and observation. The results showed that there was an accuracy rate of 88.89% with a precision value of 87.50% for restock predictions and a 100% precision value on no restock predictions. then for the recall class it is restocked 100% and for the recall class it is not restocked 50%.

Keyword :Decision Tree Algorithm, Mobile Sales, and Data Mining.

PENDAHULUAN

Saat ini perkembangan media teknologi komunikasi di Indonesia semakin berkembang dan canggih didalam kehidupan masyarakat dan itu tidak bisa dihindarkan. Sebagaimana bertambah banyak nya orang yang menggunakan handphone sebagai media komunikasi. Indonesia adalah salah satu pasar handphone yang sedang berkembang didunia, penggunaan handphone atau telepon genggam menjadi kebutuhan yang paling utama untuk bisa berkomunikasi masyarakat. saat ini penggunaan handphone dapat melakukan banyak hal, seperti melakukan kegiatan yang berkaitan dengan kehidupan sehari hari hingga banyak masyarakat menggunakan handphone untuk pekerjaan.

Perusahaan kurang menganalisa penjualan barang dan stok persediaan barang di dalam gudang, dan jika pihak perusahaan tidak pernah melakukan analisa penjualan barangnya tidak akan mengetahui apakah produk penjualan bulan ini lebih banyak dari bulan sebelumnya atau lebih sedikit dan tidak mengetahui produk apa saja yang paling laris.

Peneliti akan menggunakan data mining dan menggunakan sebuah algoritma decision tree (algoritma C4.5). Data mining diharapkan bisa memberikan solusi keputusan yang akurat didunia bisnis yang bertujuan untuk mengembangkan bisnis perusahaan dan menganalisa data untuk hubungan baru yang tidak diharapkan, tujuan melakukan penelitian ini untuk mengetahui produk apa yang paling diminati konsumen, menganalisa penjualan barang dan stok persediaan barang

didalam gudang. Manfaatnya untuk menganalisa data dan memberikan informasi data penjualan yang belum diketahui.

Menurut(Mardi, 2017) Data mining adalah sebuah proses mencari informasi yang menarik didalam data yang dipilih menggunakan teknik atau sebuah metode tertentu. Menurut (Elisa, 2017)Data mining adalah sebuah proses penggalian dan penambangan pengetahuan dari sejumlah besar data, database atau tempat penyimpanan database lainnya. Data mining ini mempunyai tujuan untuk menemukan informasi baru yang tidak diketahui dari sebuah database.

Menurut(Dhika & Destiwati, 2018) Algoritma C4.5 merupakan algoritma klasifikasi menggunakan teknik Decision tree yang sangat dikenal dan disukai karena algoritma ini memiliki kelebihan. Misalnya, dapat memproses data numerik (kontinu) dan diskrit, bisa menganalisis nilai-nilai atribut yang telah hilang, dan menghasilkan aturan yang cukup mudah untuk diinterpretasikan dan sangat cepat jika dibandingkan dari algoritma yang lainnya. Akurasi dalam memprediksi adalah kemampuan model untuk bisa memprediksi label kelas untuk data baru atau data yang sebelumnya tidak diketahui.

PENELITIAN RELEVAN

Penelitian yang dilakukan Harry Dhika, Fitriana Destiwati (2018) yang berjudul Penerapan Algoritma Data Mining Terhadap Penerima Kredit Motor yang bertujuan untuk memudahkan kinerja pada khususnya dalam memprediksi keputusan yang calon atau penerima konsumen berhak mendapatkan penolakan atau persetujuan dalam kredit yang akan diajukan.(Dhika & Destiwati, 2018)

Dalam Penelitian yang berjudul Penerapan Algoritma C4.5 untuk penentuan ketersediaan barang E-commerce yang bertujuan untuk menjaga stok yang tersedia dalam gudang menjadi stabil.(Lukhayu Pritalia, 2018)

Dalam Penelitian yang berjudul Metode penelitian kualitatif studi pustaka dan studi lapangan yang bertujuan untuk membahas metode penelitian kualitatif studi pustaka dan studi lapangan.(Darmalaksana, 2020)

Dalam Penelitian yang berjudul Analisa dan penerapan algoritma C4.5 dalam data mining untuk mengidentifikasi faktor-faktor penyebab kecelakaan kerja konstruksi PT.Arupadhatu Adisesanti yang bertujuan untuk mengetahui faktor-faktor kecelakaan kerja konstruksi.(Elisa, 2017)

Dalam Penelitian yang berjudul implementasi algoritma C4.5 untuk klasifikasi tingkat kepuasan pembeli online shop yang bertujuan untuk membuat model data mining dan mencari tingkat akurasi data.(Dimas Bayu Febriyanto, 2018)

METODE PENELITIAN

Menurut (Darmalaksana, 2020)Metode penelitian yaitu rencana pelaksanaan penelitian yang ingin dijalankan. Metode menurut (Bhalla Prem, 2017)) Metode merupakan suatu cara yang dapat digunakan untuk mencapai tujuan, sedangkan penelitian merupakan sarana untuk mencari kebenaran. Pada dasarnya penelitian merupakan dorongan untuk mengumpulkan informasi yang akan diteliti. Dalam penelitian ini, pencipta memilih strategi pengujian yang berbeda, teknik yang layak untuk mengdeskripsikan atau menggambarkan keajaiban yang ada, baik keajaiban biasa dan keajaiban yang dirancang. Sedangkan menurut (Nana & Elin, 2018)metode penelitian adalahcara ilmiah untuk memperoleh data yang valid, dengan tujuan untuk menemukan, mengembangkan, dan membuktikan suatu pengetahuan yang tertentu sehingga pada saat gilirannya dapat digunakan untuk memecahkan masalah, memahami masalah, dan mengantisipasi masalah.

HASIL DAN PEMBAHASAN

Data penjualan *handphone* ini memiliki format sebagai berikut:

- a) YA
- b) TIDAK

Dari data tersebut jadi kita mengetahui bahwa Ya artinya produk tersebut laris maka harus di restok Tidak artinya produk ini kurang laris maka tidak perlu melakukan restok karena stok produk tersebut masih tersedia. Berikut adalah perhitungan manual klasifikasi data penjualan *handphone* dalam menentukan produk yang harus di restok dan yang tidak menggunakan algoritma C4.5 yang menggunakan 18 sampel data penjualan bulan lalu sebagai data *training*.

Perancangan dilakukan menggunakan aplikasi *Rapidminer* versi 9.9.2.0

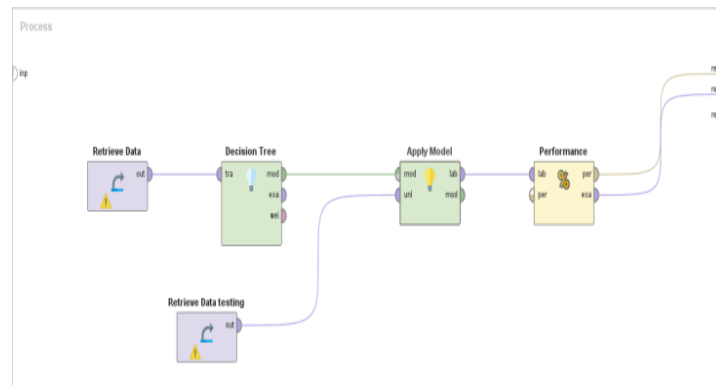
Tabel 1. Data *Training*

SPEKIFIKASI	MALL	PRODUK	PEMINAT	RESTOK
BAGUS	MBH	BARU	BANYAK	YA
BAGUS	MBH	BARU	BANYAK	YA
BAGUS	MBH	BARU	BANYAK	YA
LUMAYAN	MBH	BARU	BANYAK	YA
LUMAYAN	MBH	LAMA	BANYAK	YA
LUMAYAN	MBH	LAMA	LUMAYAN	TIDAK
BAGUS	BCP	BARU	BANYAK	YA
BAGUS	BCP	BARU	LUMAYAN	TIDAK
BAGUS	BCP	BARU	LUMAYAN	YA
LUMAYAN	BCP	BARU	LUMAYAN	YA
LUMAYAN	BCP	LAMA	LUMAYAN	TIDAK
LUMAYAN	BCP	LAMA	BANYAK	YA
BAGUS	POGE	BARU	LUMAYAN	YA
BAGUS	POGE	BARU	LUMAYAN	YA
BAGUS	POGE	BARU	LUMAYAN	TIDAK
LUMAYAN	POGE	BARU	LUMAYAN	TIDAK
LUMAYAN	POGE	LAMA	LUMAYAN	TIDAK
LUMAYAN	POGE	LAMA	BANYAK	YA

Tabel 2. Data *Testing*

SPEKIFIKASI	MALL	PRODUK	PEMINAT	RESTOK
BAGUS	MBH	BARU	BANYAK	YA
LUMAYAN	MBH	BARU	BANYAK	YA
LUMAYAN	MBH	LAMA	BANYAK	YA
BAGUS	POGE	BARU	LUMAYAN	YA
BAGUS	POGE	BARU	LUMAYAN	TIDAK
BAGUS	BCP	BARU	BANYAK	YA
BAGUS	BCP	BARU	LUMAYAN	YA
LUMAYAN	BCP	LAMA	LUMAYAN	TIDAK
LUMAYAN	POGE	LAMA	BANYAK	YA

Data yang akan digunakan dalam aplikasi digunakan data *training*, dari seluruh data dipisah menjadi dua bagian yaitu data *training* dan data *testing* seperti tabel 1 dan 2, dari meta data tersebut dilakukan pengolahan data pada *RapidMiner* versi 9.9.2.0 dengan *mainproses* sebagai berikut:



Gambar 1. Main Proses

Mainproses dilakukan dengan menambahkan data *fileexcel*, dan memasukkan ke dalam *boxdecisiontree* untuk menghasilkan pohon keputusan dari data tersebut. Lalu dimasukkan ke *boxapplymodel* untuk menghasilkan model dari data tersebut. Dan memasukkan *boxperformance* untuk dapat menghasilkan pengukuran yang maksimal. Kemudian diperoleh tingkat pengukuran akurasi sebagai berikut:

accuracy: 88.89%

	true YA	true TIDAK	class precision
pred. YA	7	1	87.50%
pred. TIDAK	0	1	100.00%
class recall	100.00%	50.00%	

Gambar 2. Pengukuran Akurasi Data

Hasil pengukuran akurasi data diperoleh bahwa keseluruhan nilainya mencapai 88,89% dari gambar 4.3 pengukuran akurasi data diketahui prediksi produk yang harus di restok dengan *true* YA mencapai nilai 7 produk dan untuk *true* TIDAK sebanyak 1 produk dengan hasil capaian nilai presisi mencapai sebesar 87,50% sedangkan untuk prediksi produk TIDAK restok untuk *true* YA mencapai nilai 0 produk dan untuk *true* TIDAK terdapat 1 produk, *persentase* TIDAK adalah sebesar 100% produk. Untuk *recall* data produk sangat tinggi yaitu untuk *classrecalltrue* YA mencapai 100% sedangkan untuk *classrecalltrue* TIDAK mencapai nilai 50% produk. Pengukur akurasi melihat data lebih besar pada produk di restok dengan tingkat akurasi tertinggi yang berbanding sedikit dengan pengukuran *precision*. Pengukuran berikutnya sebagai pembanding yakni *precision* data sebagai berikut:

precision: 100.00% (positive class: TIDAK)

	true YA	true TIDAK	class precision
pred. YA	7	1	87.50%
pred. TIDAK	0	1	100.00%
class recall	100.00%	50.00%	

Gambar 3. Pengukuran Precision Data

Hasil dari pengukuran *precision* data sudah diperoleh bahwa secara keseluruhan dari nilainya mencapai 100% hasil tersebut sama seperti hasil dari pengukuran data akurasi dengan hasil tabel tersebut dapat diketahui prediksi YA produk yang restok dengan *true* YA mencapai 7 produk dan untuk *true* TIDAK sebanyak 1 produk dengan hasil capaian nilai presisi mencapai sebesar 87,50% sedangkan untuk prediksi produk TIDAK restok untuk *true* YA mencapai nilai 0 produk dan untuk

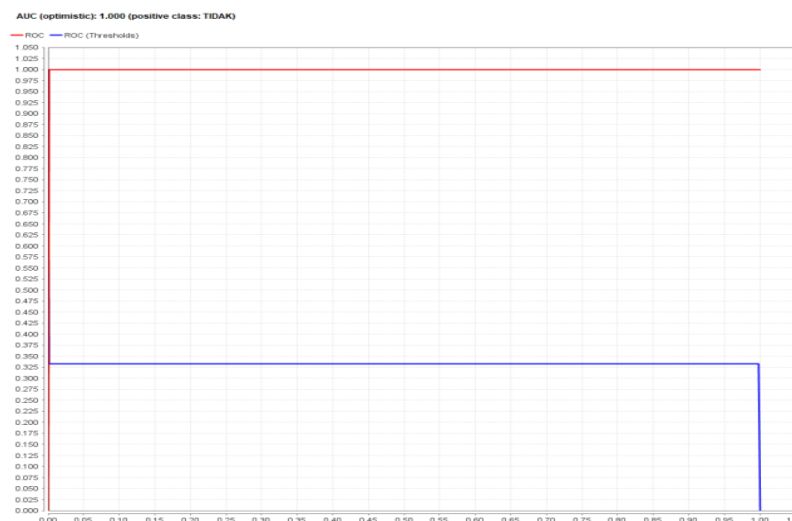
true TIDAK terdapat 1 produk, *persentase* TIDAK restok adalah sebesar 100% produk. Untuk *recall* data produk sangatlah tinggi yakni untuk nilai *classrecalltrue* YA mencapai 100% sedangkan untuk nilai *classrecalltrue* TIDAK mencapai nilai 50% dalam hal ini maka data difokuskan pada kebalikan dari hasil *accuracy*. Nilai dari pengukuran *confusionmatrix* selanjutnya yakni dari *classrecall* yaitu sebagai berikut:

recall: 50.00% (positive class: TIDAK)

	true YA	true TIDAK	class precision
pred. YA	7	1	87.50%
pred. TIDAK	0	1	100.00%
class recall	100.00%	50.00%	

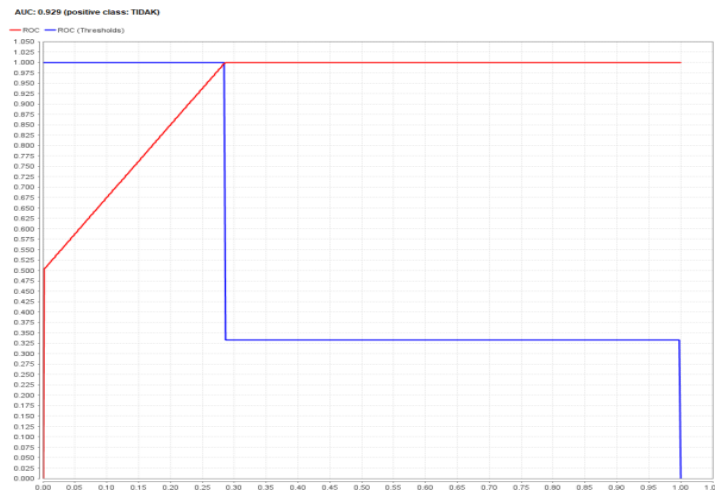
Gambar 4. Pengukuran Recall Data

Hasil dari pengukuran *recall* data jika diperhatikan untuk prediksi TIDAK restok dan YA restok adalah sama. Yang membedakan adalah nilai dari *recall* sebesar 50% yang membedakan dengan lainnya. Nilainya mengartikan bahwa untuk nilai atau hasil keputusan dalam produk yang harus direstok sangat sedikit sekali yang TIDAK restok sehingga yang merupakan *positiveclass* adalah “TIDAK”. Setelah semua tahapan evaluasi untuk *confusionmatrix* dilakukan, maka akan dilakukan analisis evaluasi komparatif yaitu pengukuran *receiveroperatingcharacter* (ROC) sebagai berikut:



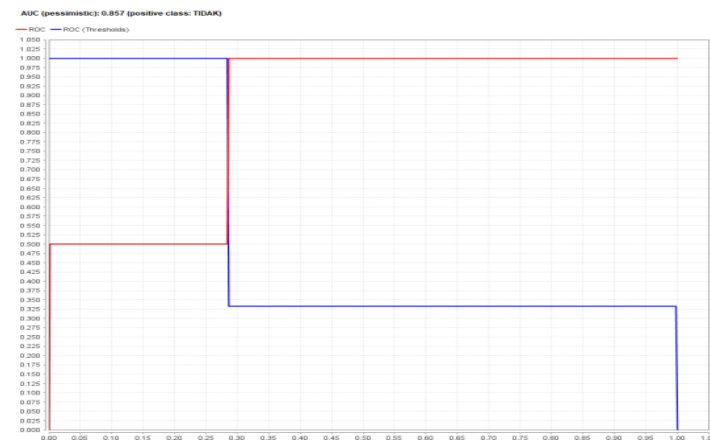
Gambar 5. Pengukuran UAC Optimistic

Nilai ROC dilakukan pengukuran terlebih dahulu terhadap *AreaUnderCurvenya* (AUC) *optimistic* dengan nilai sebesar 1,000 terhadap *class* TIDAK dan untuk *optimistic* dengan nilai sebesar 0,325 terhadap *class* YA . Pengukuran dilakukan berikutnya terhadap AUC normal yakni:



Gambar 6. Pengukuran AUC

Nilai yang diperoleh dari pengukuran AUC yakni sebesar 0,929 untuk *class* TIDAK. Dengan nilai sebesar 0,500 terhadap *class* TIDAK dan nilai sebesar 1,000 terhadap *class* YA. Selanjutnya dilakukan pengukuran AUC untuk *pessimistic*-nya dilihat pada gambar 7 sebagai berikut:



Gambar 7. Pengukuran AUC Pessimistic

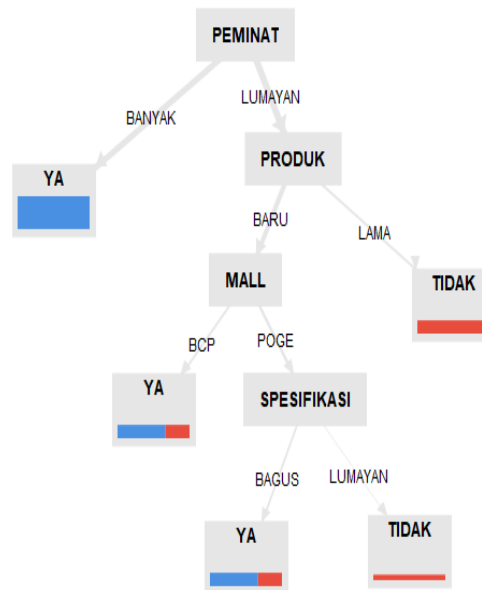
Nilai yang diperoleh dari AUC *pessimistic* yakni 0,857 untuk *classpredictive* TIDAK. Pengukuran berikutnya dilakukan untuk mengetahui *performance* terhadap *vector* yang diperoleh dan dari hasil keseluruhan maka diperoleh pohon keputusannya atau *decisiontree* nya sebagai berikut:

PerformanceVector

```

PerformanceVector:
accuracy: 88.89%
ConfusionMatrix:
True:  YA    TIDAK
YA:    7     1
TIDAK: 0     1
precision: 100.00% (positive class: TIDAK)
ConfusionMatrix:
True:  YA    TIDAK
YA:    7     1
TIDAK: 0     1
recall: 50.00% (positive class: TIDAK)
ConfusionMatrix:
True:  YA    TIDAK
YA:    7     1
TIDAK: 0     1
AUC (optimistic): 1.000 (positive class: TIDAK)
AUC: 0.929 (positive class: TIDAK)
AUC (pessimistic): 0.857 (positive class: TIDAK)
    
```

Gambar 8. Performance Vector



Gambar 9. Pohon keputusan

Dari hasil gambar 4.10 pohon algoritma produk yang harus direstok lebih jelasnya itu jika Peminat dari produk tersebut banyak maka barang laris dan harus direstok, jika peminat lumayan maka dilihat dari produk tersebut, apakah produk baru atau lama. Jika produk lama maka barang kurang laris dan stok masih tersedia jadi tidak perlu direstok, sedangkan jika produk baru dilihat dari tempat penjualannya atau mall, apakah mall BCP atau mall POGE. Jika mall BCP maka barang laris dan harus melakukan restok, sedangkan jika mall POGE dilihat dari Spesifikasi produk tersebut, apakah bagus atau lumayan bagus. Jika Spesifikasi produk tersebut bagus maka barang laris dan harus melakukan restok, sedangkan jika spesifikasi produk tersebut lumayan bagus maka produk tersebut kurang laris dan produk masih tersedia dan tidak perlu melakukan restok. Pada gambar 4.10 pohon algoritma C4.5 Produk yang harus restok jika dijabarkan kedalam bahasa yang lebih sederhana dibuat dalam bentuk pohon *text* maka akan diperoleh hasil sebagai berikut:

```

TOTAL > 93.500
|  PRODUK = BARU
|  |  TOTAL > 109.500: YA {YA=7, TIDAK=0}
|  |  TOTAL ≤ 109.500
|  |  |  MALL = BCP: TIDAK {YA=0, TIDAK=1}
|  |  |  MALL = POGE: YA {YA=2, TIDAK=1}
|  PRODUK = LAMA
|  |  Merek = Vivo Y12: YA {YA=2, TIDAK=1}
|  |  Merek = Vivo Y15: TIDAK {YA=0, TIDAK=3}
TOTAL ≤ 93.500: TIDAK {YA=0, TIDAK=1}
  
```

Gambar 10. Performance Vector

Untuk hasil pengukuran tingkat akurasi dalam data *mining*, nilai-nilai tersebut dapat dibagi menjadi beberapa kelompok. (Gorunescu, 2011) Yaitu:

- a) 0.90-1.00 = klasifikasi golongan sangat baik
- b) 0.80-0.90 = klasifikasi golongan baik
- c) 0.70-0.80 = klasifikasi golongan cukup
- d) 0.60-0.70 = klasifikasi golongan buruk
- e) 0.50-0.60 = klasifikasi golongan salah

Dari nilai hasil pengukuran akurasi maka dapat disimpulkan sesuai dengan pengelompokan klasifikasi diatas, masuk pada klasifikasi golongan cukup untuk penerapan algoritma C4.5 pada penjualan *handphone* yaitu dengan menentukan produk mana yang harus direstok dan mana yang tidak direstok.

SIMPULAN

Permasalahan pada penjualan *handphone* yaitu menentukan produk mana yang laris dan tidak laris serta produk mana yang harus direstok dan yang tidak dapat diselesaikan menggunakan teknik data *mining*, yaitu dengan algoritma C4.5. dengan melakukan metode ini akan menghasilkan tingkat akurasi dengan *persentase* sebesar 88,89%.

DAFTAR PUSTAKA

- [1] Bhalla Prem. (2017). Tatacara Ritual dan Tradisi Hindu. (Surabaya: Paramita, 2010). *Paramita*, 3(1), 78–95.
- [2] Darmalaksana, W. (2020). Metode Penelitian Kualitatif Studi Pustaka dan Studi Lapangan. *Pre-Print Digital Library UIN Sunan Gunung Djati Bandung*, 1–6. [http://digilib.uinsgd.ac.id/32855/1/Metode Penelitian Kualitatif.pdf](http://digilib.uinsgd.ac.id/32855/1/Metode%20Penelitian%20Kualitatif.pdf)
- [3] Dhika, H., & Destiawati, F. (2018). Penerapan Algoritma Data Mining Terhadap Penerima Kredit Motor. 1–11. <https://doi.org/10.31227/osf.io/g8vy5>
- [4] Dimas Bayu Febriyanto. (2018). Implementasi Algoritma C4 . 5 Untuk Klasifikasi Tingkat Kepuasan Pembeli Online Shop. *Jurnal Riset Komputer*, 5(6), 569–575.
- [5] Elisa, E. (2017). Analisa dan Penerapan Algoritma C4.5 Dalam Data Mining Untuk Mengidentifikasi Faktor-Faktor Penyebab Kecelakaan Kerja Kontruksi PT.Arupadhatu Adisesanti. *Jurnal Online Informatika*, 2(1), 36. <https://doi.org/10.15575/join.v2i1.71>
- [6] Gorunescu, F. (2011). *Data Mining Concept Model and Techniques*.
- [7] Lukhayu Pritalia, G. (2018). Penerapan Algoritma C4.5 untuk Penentuan Ketersediaan Barang E-commerce. *Indonesian Journal of Information Systems*, 1(1), 47–56. <https://doi.org/10.24002/ijis.v1i1.1727>
- [7] Mardi, Y. (2017). Data Mining : Klasifikasi Menggunakan Algoritma C4.5. *Edik Informatika*, 2(2), 213–219. <https://doi.org/10.22202/ei.2016.v2i2.1465>
- [8] Marudut Mulia Siregar, V. (2017). Perancangan Aplikasi Data Mining Untuk Memprediksi Penjualan Menggunakan Metode Decision Tree Pada Apotik Ths Pematangsiantar. *Jurnal Murni Sadar*, 7(1), 51–61.
- [9] Nana, D., & Elin, H. (2018). Memilih Metode Penelitian Yang Tepat: Bagi Penelitian Bidang Ilmu Manajemen. *Jurnal Ilmu Manajemen*, 5(1), 288. <https://jurnal.unigal.ac.id/index.php/ekonomi/article/view/1359>