

Studi Literature Analisis Potensi Pasar Marketplace terhadap Penjualan

Sya'Roni ¹, Cahyo Crysdiandian ²

Universitas Islam Negeri Maulana Malik Ibrahim Malang, Indonesia

Article Info

Article History

Received : 28-11-2022

Revises : 13-12-2022

Accepted : 02-01-2023

Keywords

KNN

C4.5

Marketplace

Produk

✉ Corresponding Author

Sya'Roni,

Universitas Islam Negeri

Maulana Malik Ibrahim

Malang,

Tel. +62 85790941311

200605220006@student.uin-

malang.ac.id

ABSTRACT

Marketplace allows Customer to Customer (C2C) transactions between consumers without being bound by place and time. This change also occurs in human spending habits. So it becomes an opportunity for sellers to market their wares. In facing market competition, business people also need analysis to find out what products are selling best. However, there are many factors that affect the complexity of the marketplace in Indonesia. Then a classification-based simulation using KNN and C4.5 is needed, where the weight of each sales product group affects market potential so that it benefits sellers to choose which marketplace is suitable for the goods to be sold. So it can be concluded that (1). Factors that influence the complexity of marketplaces in Indonesia are price, number of sales, discounts, ratings and reviews. (2). The most optimal method used for analysis of sales market potential is K-Nearest Neighbor and C4.5.

PENDAHULUAN

Salah satu cara untuk memenuhi kebutuhan manusia adalah dengan melakukan kegiatan ekonomi. Perdagangan merupakan salah satu jenis usaha yang dilakukan oleh banyak orang di seluruh dunia. Bisa apa saja mulai dari berdagang berbagai kebutuhan primer atau sehari-hari hingga menjual barang yang sangat mahal demi kepuasan manusia.

Berbeda dengan dulu, ketika penjual harus memiliki tempat atau toko untuk berjualan, saat ini penjual tidak perlu memiliki tempat atau toko untuk berjualan. Sebaliknya, penjual bisa berjualan secara *online* hanya dengan ponsel atau laptop, koneksi internet, dan barang yang akan dijual. Setelah penjual memiliki semua yang kita butuhkan, yang harus kita lakukan untuk menjual produk kita adalah mengunjungi ke beberapa *marketplace*. penjual juga mendapatkan banyak keuntungan, seperti tidak perlu menyewa tempat atau toko, mendapat notifikasi ketika pelanggan hendak membeli produk yang kita jual, menerima rangkuman data transaksi penjualan pasar yang cukup akurat, dan masih banyak lagi.

Kegiatan belanja online termasuk dalam industri *e-commerce*. Industri *e-commerce* telah tumbuh dengan mantap sejak awal, dan diperkirakan tren ini akan berlanjut hingga tahun 2020 [1]. Perluasan pasar *e-commerce* menunjukkan bahwa *e-commerce* menghadirkan skenario bisnis yang cukup menjanjikan. Seiring dengan perkembangannya, mulai bermunculan berbagai pusat niaga yang penting bagi bisnis berbasis web yang juga

memberikan tempat kepada konsumen untuk melakukan berbagai transaksi secara *online*, salah satunya adalah belanja *online*.

Setiap bisnis dalam lingkungan bisnis global saat ini harus mampu bersaing dengan meningkatkan kinerja dan menyajikan laporan yang tepat agar bisnis dapat mengambil keputusan berdasarkan hasil kinerja dan laporan. Pengolahan data, di sisi lain dipengaruhi oleh perkembangan teknologi informasi untuk menghasilkan informasi yang lebih cepat, lebih akurat, terkini, dan relevan untuk pengambilan keputusan.

Marketplace memungkinkan transaksi antara konsumen dengan konsumen atau *Customer to Customer (C2C)* tanpa terikat tempat dan waktu. Data *iprice* pada kuartal empat tahun 2021 perkembangan data pengguna *marketplace* di Indonesia sudah mencapai 382 juta pengguna, dimana penjual masih bingung di *marketplace* mana barang dagangannya akan dijual, karena setiap *marketplace* mempunyai potensi pasarnya masing-masing sehingga membutuhkan klasifikasi potensi pasar untuk mengetahui potensi setiap *marketplace* [1]. Salah satu tujuan yang banyak dihasilkan adalah klasifikasi. Tujuan klasifikasi adalah untuk memperkirakan kelas label objek yang tidak diketahui atau untuk membagi sesuatu ke dalam kelasnya menggunakan fungsi pengelompokan atau pengelompokan yang menjelaskan atau membedakan konsep atau kelas data.

Karena hampir semua toko baik *online* maupun *offline* saat ini menyimpan data penjualan dalam sistem yang terkomputerisasi, maka akan dihasilkan banyak data transaksi. *Volume record* transaksi penjualan yang tersimpan di data transaksi *database* penjualan barang sangat tinggi, sehingga mengakibatkan jumlah data yang terus bertambah. Menginvestigasi akumulasi data untuk menemukan pola penjualan produk yang dapat digunakan untuk menganalisis pasar dan meramalkan penjualan di masa depan. Menggunakan metode atau teknik untuk memecahkan masalah klasifikasi bertujuan untuk mempermudah proses klasifikasi .

Dalam menghadapi persaingan pasar, Para pelaku bisnis juga membutuhkan analisis untuk mengetahui produk apa yang paling laku [2]. Sedangkan algoritma *K-Nearest Neighbor* [3] melakukan Klasifikasi Produk pada toko online yang telah diuji pada tiga kategori produk. Dengan menggunakan nilai $k = 3$, rata-rata nilai *accuracy*, *precision*, dan *recall* paling rendah saat melakukan klasifikasi produk. Kemudian, dengan menggunakan $k = 5$ atau $k = 7$ menghasilkan nilai *recall*, *presisi*, dan akurasi tertinggi. Karena memiliki nilai tinggi nilai *precision*, dan *recall* selain nilai *accuracy* yang tinggi, algoritma *K-Nearest Neighbor* dapat digunakan secara efektif. Karena jarak antara setiap data uji dan seluruh data pelatihan perlu dihitung, proses klasifikasi semakin lama semakin banyak dataset yang digunakan.

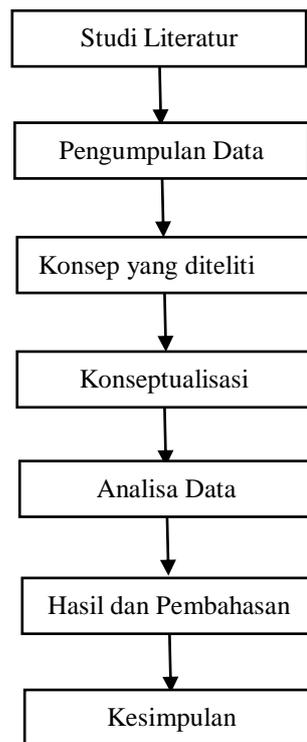
Pembeli dapat diprediksi dengan cukup akurat menggunakan perhitungan *C4.5* [4]. Efek samping dari perkiraan ketepatan informasi yang diperoleh dari penyusunannya dengan nilai tambah *Confusion Matrix* adalah sebesar 91,33%. Kemudian tingkat ketepatan informasi pengujian mencapai 70,00%.

Berdasarkan latar belakang permasalahan diatas, rumusan permasalahan yang diangkat pada penelitian ini adalah penerapan *K-Nearest Neighbor* dan algoritma *C4.5* untuk melakukan klasifikasi data produk dari beberapa *marketplace* di Indonesia. Penulis mencoba menerapkan teknik data mining metode *K-Nearest Neighbor* dan algoritma *C4.5* pada analisis potensi pasar *marketplace* terhadap penjualan dan diharapkan dapat memberikan informasi berupa klasifikasi penjualan produk (laris dan tidak laris) dimana model *K-Nearest Neighbor* dan algoritma *C4.5* masing-masing mempunyai jarak dan pohon keputusan yang mampu memberikan akurasi klasifikasi yang baik. Sehingga kedepannya penjual dapat melakukan analisa potensi pasar *marketplace* terhadap penjualannya.

METODE

Tujuan dari penelitian ini adalah untuk mengeksplorasi berbagai studi berbasis klasifikasi pada Analisa potensi pasar dalam hal parameter akurasi, nilai atribut tertinggi, kesenjangan

kompetensi, dan peringkat atribut. Berikut Langkah-langkah untuk melakukan studi literatur seperti terlihat Gambar 1.



Gambar 1. Alur Studi Literatur.

Berdasarkan pada Gambar 1. dapat dijelaskan tentang langkah-langkah penelitian ini:

- **Pengumpulan Data**
Penelitian ini menggunakan data dari jurnal-jurnal nasional dan internasional. Memanfaatkan Mendeley.com dan *search engine google scholar* adalah cara terbaik untuk menemukan jurnal yang terkait dengan metode klasifikasi dan sistem rekomendasi. Kriteria tahun terbit jurnal tidak lebih dari lima tahun dari 2017 hingga 2022 yang dapat diakses full text dalam format pdf dan bersifat ilmiah digunakan dalam proses pemilihan data.
- **Konsep yang diteliti**
Melakukan eksplorasi jurnal-jurnal yang sudah terkumpul
- **Konseptualisasi**
dengan memahami teknik klasifikasi karya peneliti lain pada subjek penelitian. Pada tahap ini, jurnal dengan hasil pengujian terbaik menggunakan metode klasifikasi dipilih dari jurnal yang terkumpul dan dibandingkan satu sama lain untuk menghasilkan jurnal dengan hasil pengujian terbaik.
- **Analisa**
Pada tahap ini analisis data dilakukan dengan mengorganisasikan, menyusun data secara berurutan, dan menyusun data sesuai dengan kategori data. Selain itu, visualisasi data dilakukan dengan menyajikan hasil analisis data dalam bentuk tabel untuk membantu peneliti dalam memahami temuan.
- **Hasil dan Pembahasan**
Pada tahap ini, hasil pengujian dari masing-masing jurnal yang terkumpul diperiksa untuk membuat perbandingan jurnal ke jurnal. Agar memudahkan memahami masing-masing jurnal dibuatlah kerangka teori dari penelitian ini.
- **Kesimpulan**
Tahap ini menghasilkan pengujian terbaik dari metode masing-masing jurnal.

HASIL DAN PEMBAHASAN

• Analisa Metode

Penelitian terkait mengenai pembahasan analisis potensi pasar *marketplace* terhadap penjualan telah banyak dilakukan pada penelitian sebelumnya.

Menghadapi persaingan pasar, pelaku bisnis juga harus memiliki akses ke analisis yang dapat memprediksi produk yang paling populer [2]. Karena *Naive Bayes* didasarkan pada asumsi penyederhanaan bahwa nilai atribut independen secara kondisional ketika diberi nilai output, penulis menggunakan metode *Naive Bayes*. Pemanfaatan *Naive Bayes* dalam ulasan ini memerlukan jumlah informasi penyiapan yang cukup terbatas untuk menentukan pengukur batas yang diperlukan dalam sistem pengelompokan. Data string bersifat konstan dan data numerik kontinu dibedakan dalam metode *Naive Bayes*. Saat menentukan perbedaan ini akan menentukan.

Pembobotan algoritma *K-Nearest Neighbor* menggunakan metode TF-IDF untuk menentukan nilai akurasi [3]. Dengan menggunakan rumus *cosine similarity*, jarak kemiripan antara dokumen uji dan dokumen training dihitung menggunakan hasil pembobotan TF-IDF. Algoritma *K-Nearest Neighbor* digunakan untuk mengklasifikasikan dokumen setelah diperoleh hasil *cosine similarity*. Selain itu, hasil menunjukkan bahwa algoritma *K-Nearest Neighbor* efektif sehingga memungkinkan untuk rekomendasinya.

K-Modes Clustering dengan nilai $k = 2,3,4,\dots,10$, digunakan untuk mengklasifikasikan tingkat minat belanja online [5]. Jumlah *cluster* terbaik ditentukan dengan menggunakan *Davies-Boulden Index* (DBI). Jumlah *cluster* terbaik diperoleh dari analisis, dengan $k = 9$ dan nilai DBI 1,3427. Laki-laki muda berusia antara 15 dan 24 tahun membentuk *cluster* ini. Anggota *Cluster 5* adalah *cluster* terbaik, memiliki minat yang besar belanja di media sosial dan *marketplace*. Shopee, Bukalapak, dan Tokopedia adalah *marketplace* yang sangat diminati. Instagram, Facebook, dan Media Chatting adalah platform media sosial yang sangat diminati.

Penggunaan algoritma *K-Means* untuk mengelompokkan data menjadi beberapa *cluster* berdasarkan kesamaannya satu sama lain [6]. Ia melakukan sembilan iterasi, dengan dua *cluster*, C0 dan C1, dimana C0 merepresentasikan produk yang tidak laku dan C1 merepresentasikan produk yang sangat laris. Pengelompokan ini dapat dilihat pada nilai *centroid* iterasi kedelapan dan kesembilan, dimana nilai *centroid* tersebut berasal dari rata-rata nilai penjualan bulan februari. laris, dan kelompok C1 memiliki produk yang paling. Kelompok produk C0 ada 657 barang yang tidak laris, sedangkan kelompok produk C1 ada 70 barang yang laris.

Melakukan pemilihan *e-marketplace* menggunakan teknik pengambilan keputusan metode *fuzzy*, dengan kualitas layanan, kualitas sistem, kualitas sistem informasi dan kualitas penyedia layanan sebagai kriteria [7]. TK, BL, dan SH adalah tiga alternatif pilihan. Urutan nilai integral total untuk setiap derajat optimisme adalah $SH > TP > BL$ berdasarkan hasil perhitungan.

Pembeli dapat mengantisipasi dengan cukup akurat menggunakan perhitungan *C4.5* efek samping memperkirakan presisi informasi yang didapat dari penyiapan informasi dengan nilai tambah *Confusion Matrix* adalah 91,33% [4]. Kemudian tingkat presisi informasi pengujian mencapai 70,00%. Jika dilihat dengan ROC, data testing memiliki presisi 0,838 dan memiliki akurasi 0,750.

Untuk mencari probabilita dari masing-masing kelas, membuat kalkulasi menggunakan algoritma *naive bayes* untuk memutuskan kelas mana yang “Laku”, “Sedang” dan “Tidak Laku” dalam data pelatihan [8]. Untuk menghitung probabilitas akhir setiap kelas, data pelatihan harus dikonversi ke nilai yang telah ditentukan. Nilai probabilitas untuk setiap atribut dikalikan dengan nilai probabilitas kelas untuk setiap atribut agar dapat

membandingkan nilai tertinggi berdasarkan hasil yang telah ditentukan untuk setiap kelas. Hasilnya adalah jika kelas “TIDAK LAKU” bernilai lebih besar maka hasilnya “TIDAK LAKU”. Begitu pula dengan “SEDANG” dan “LAKU”.

Penelitiannya menggunakan algoritma CHAID untuk harga, kota penjual, jenis produk, jumlah penonton, dan peringkat adalah variabel independen. Selain itu, lebih baik menggunakan variabel independen sebagai simpul pemisah [9]. Keputusan dibuat dengan membandingkan nilai p dari setiap variabel independen dari proses penggabungan dibedakan sebagai variabel utama dalam membentuk pohon keputusan dan menjadikan jumlah *viewer* sebagai elemen yang paling memengaruhi transaksi di Tokopedia.

Algoritma *K-Means* digunakan untuk menentukan jarak data dari setiap *centroid* [10]. Perhitungan nilai *cluster* dihasilkan dari perhitungan ini, yang menggunakan rumus jarak *Euclidean*. Hasil analisis penelitian ini membawa kita pada kesimpulan bahwa calon pelanggan SPS Motor Honda Lombok Timur adalah orang-orang berusia antara 33 sampai 46 tahun yang berprofesi sebagai petani dan wiraswasta. Hal ini karena nasabah tersebut merupakan usia kerja dan berpenghasilan.

Penerapan FP-Growth pada data penjualan, dapat ditarik kesimpulan bahwa data penjualan dapat diolah menggunakan teknik data mining untuk mengidentifikasi pola dalam *frequent itemset* dan menggunakannya sebagai informasi penting saat memilih strategi pemasaran dan pengaturan stok [11]. Selain itu, algoritma FP-Growth memiliki metode yang sangat efektif untuk mengidentifikasi pola penjualan produk.

Pendekatan sederhana, terukur, dan non-parametrik untuk klasifikasi teks pendek dengan memilih kata-kata indikatif topik sebagai kata-kata kueri dari segmen teks pendek untuk diklasifikasikan dan mengambil segmen teks berlabel yang paling cocok dengan kata-kata kueri [12]. Label kelas yang diprediksi adalah suara terbanyak dari hasil pencarian. Hasilnya *Fuzzy* akurasi mencapai 81.27% dan *K-Nearest Neighbor* mencapai 85.17%.

Data uji dibagi menjadi dua bagian yakni, setengah digunakan untuk merujuk data riwayat pengguna, dan disebut sebagai data referensi masa lalu [13]. Sedangkan sebagian lagi dari data yang digunakan untuk mengevaluasi hasil dari algoritma rekomendasi, disebut sebagai data prediktif. Berdasarkan riwayat pembelian dan skor produk sebelumnya, algoritme ini menghitung produk yang direkomendasikan serupa dan berbeda. Produk-produk ini diperiksa terhadap kumpulan data prediktif apakah produk yang direkomendasikan ada dalam kumpulan data prediktif. Kehadiran produk rekomendasi dalam kumpulan data prediktif memberikan kepercayaan dan cakupan produk serupa dalam daftar rekomendasi. Karena algoritme yang diusulkan juga mencakup produk yang berbeda untuk rekomendasi, keberadaan produk yang tidak diketahui dalam daftar rekomendasi untuk membandingkan kumpulan data prediktif juga dievaluasi.

Data-data seperti nama produk, harga produk, ulasan dan peringkat produk yang terkait dengan produk yang disarankan dari *Flipkart* dan *Amazon* menggunakan alat Otomatisasi Proses Robotik [14]. Data dibersihkan untuk menghilangkan nilai yang mengganggu dan berdasarkan permintaan pengguna, produk dari *Flipkart* dan *Amazon* disarankan di portal *E-Commerce* berdasarkan preferensi pengguna pengguna dapat membeli produk dari situs mana pun yang sepenuhnya bergantung pada pengguna. Alat *RPA* Melalui *UiPath* digunakan untuk membuat robot perangkat lunak untuk *Flipkart* dan *Amazon*, bot ini digunakan untuk mengikis data dari *Flipkart* dan *Amazon* dan melakukan operasi pembersihan dan membuang data ke *database* yang dapat digunakan untuk penggunaan di masa mendatang. Dan bot dapat dipicu setiap hari untuk pengumpulan data untuk pembaruan basis data. Portal *E-Commerce* ini membantu untuk mengkategorikan dan membedakan perilaku pembelian produk dan memenuhi kebutuhan pelanggan akhir.

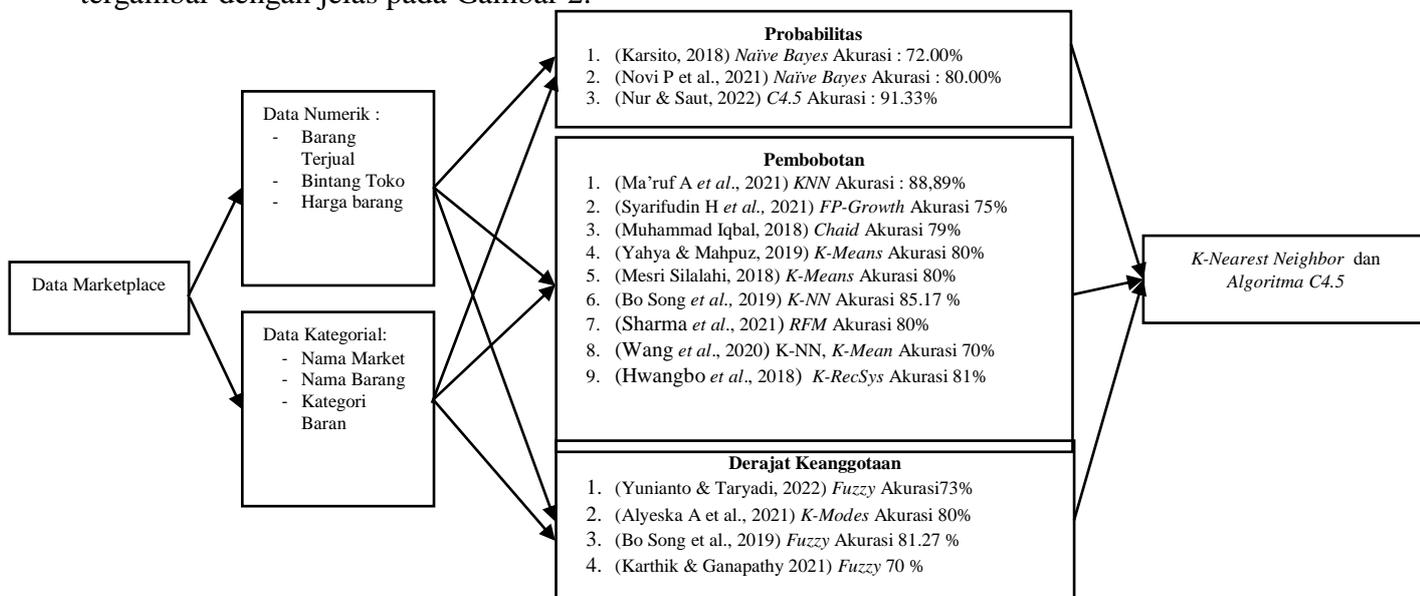
Produk e-commerce yang dipersonalisasi sistem rekomendasi berbasis pembelajaran representasi *clustering* [15]. Metode *K-Nearest Neighbor* tradisional memiliki keterbatasan dalam memilih kumpulan objek yang berdekatan. Kami memperkenalkan faktor tetangga dan

fungsi waktu dan memanfaatkan model pemilihan dinamis untuk memilih kumpulan objek yang berdekatan.

Sistem penyaringan dan rekomendasi kolaboratif dunia nyata yang digunakan oleh perusahaan mode besar Korea yang menjual produk mode secara *online* dan *offline* [16]. Karakteristik unik dari lingkungan rekomendasi perusahaan adalah sebagai berikut: Pertama, produk dijual baik di perusahaan *online* maupun toko *offline*. Kedua, item *fashion* biasanya bersifat musiman, sehingga preferensi umum pelanggan berubah seiring dengan musim. Terakhir, pelanggan biasanya membeli item untuk melengkapi atau mengganti item yang telah mereka beli sebelumnya. Kami mengusulkan sistem baru yang disebut *K-RecSys*, yang mengambil karakteristik domain yang disebutkan di atas dan memperluas algoritme penyaringan kolaboratif berbasis item standar. *K-RecSys* menggabungkan data klik produk dari penjualan *online* dan *offline* yang diberi bobot untuk mencerminkan preferensi pelanggan dari *online* dan *offline*. Ini juga menggunakan fungsi preferensi untuk mencerminkan pergeseran preferensi dari waktu ke waktu, dan menggunakan informasi tentang kategori produk untuk merekomendasikan produk terternatif dan pelengkap.

• **Kerangka Teori**

Berdasarkan hasil penelitian dari masing-masing jurnal untuk analisis literatur, maka dibuat kerangka teori untuk penentuan metode algoritma yang paling optimal seperti tergambar dengan jelas pada Gambar 2.



Gambar 2. Kerangka Teori

Berdasarkan Gambar 2. Kerangka teori untuk menentukan algoritma yang memiliki nilai performa tinggi terlihat pada Tabel 1.

Tabel 1. Daftar Jurnal

No.	Judul Jurnal	Metode Penelitian	Penulis, Tahun	Hasil Penelitian
1	Data Mining Untuk Klasifikasi Produk Menggunakan Algoritma <i>K-Nearest Neighbor</i> Pada Toko <i>Online</i>	<i>K-Nearest Neighbor</i>	(Ma'ruf Aziz Muzani et al., 2021)	Hasil : Akurasi : 88,89%
2	Prediction Tingkat Potensi Pembeli Pada Hari Special Di Toko Mitra Shopee	Algoritma C4.5	(Nur Artisya & Saut Pintubipar Saragih, 2022)	Hasil : Akurasi : 91.33%
3	Implementasi Fuzzy Decision Making Untuk Pemilihan <i>Marketplace</i>	Fuzzy Decision Making	(Era Yuniyanto & Taryadi, 2022)	Hasil : Akurasi 73 %

Studi Literature Analisis Potensi Pasar Marketplace terhadap Penjualan
Sya'Roni, Cahyo Crysdiان

4	Prediksi Potensi Penjualan Produk Delifrance Dengan Metode <i>Naive Bayes</i> Di PT. Pangan Lestari	<i>Naive Bayes</i>	(Karsito & Winda Monika Sari, 2018)	Hasil : Akurasi : 72.00%
5	Penerapan Data Mining Prediksi Penjualan Barang Elektronik Terlaris Menggunakan Algoritma Naive Bayes (Study Kasus : Planet Cash And Credit Cabang Muara Enim)	<i>Naive Bayes</i>	(Novi Pransiska et al., 2021)	Hasil : Akurasi : 80.00%
6	Analisa Metode Association Rule Menggunakan Algoritma FP-Growth Terhadap Data Penjualan (Study Kasus Toko Berkah)	Association Rule dan Algoritma FP-Growth	(Syarifudin Herdyansyah et al., 2021)	Hasil : Akurasi 75%
7	Analisis Dan Prediksi Penjualan Pada <i>Marketplace</i> Berdasarkan Pendekatan Klasifikasi Dengan Metode Pohon Keputusan (Studi Kasus Pada Data Katalog Dan Penjualan Tas Pria & Wanita Di Tokopedia Secara Nasional)	CHAID Algorithm	(Muhammad Iqbal Islami dan Adhi Prasetio, 2018)	Hasil : Akurasi 79%
8	Penggunaan Algoritma <i>K-Means</i> Untuk Menganalisis Pelanggan Potensial Pada Dealer SPS Motor Honda Lombok Timur Nusa Tenggara Barat	<i>K-Means</i>	(Yahya & Mahpuz, 2019)	Hasil : Akurasi 80%
9	Analisis <i>Clustering</i> Menggunakan Algoritma <i>K-Means</i> Terhadap Penjualan Produk Pada PT Batamas Niaga Jaya	<i>K-Means</i>	(Mesri Silalahi, 2018)	Hasil : Akurasi 80%
10	Penerapan Algoritma K-Modes <i>Clustering</i> dengan Validasi Davies Bouldin Index pada Pengelompokan Tingkat Minat Belanja <i>Online</i> di Provinsi Daerah Istimewa Yogyakarta	K-Modes <i>Clustering</i>	Alyeska Astri Az-zahra et al., 2021)	Hasil : Akurasi 80 %
11	Cross-border e-commerce commodity risk assessment using text mining and fuzzy rule-based reasoning	fuzzy rule-based reasoning and <i>K-NEAREST NEIGHBOR</i> - based classifiers	(Bo Song et al., 2019)	Hasil : Fuzzy : 81.27% <i>K-Nearest Neighbor</i> : 85.17%
12	A fuzzy recommendation system for predicting the customers interests using sentiment analysis and ontology in e-commerce	fuzzy	(Karthik and Ganapathy 2021)	Hasil : Akurasi 70%
13	E-Commerce product comparison portal for classification of customer data based on data mining	RFM	(Sharma et al. 2021)	Hasil : Akurasi 80 %
14	E-Commerce Personalized Recommendation Analysis by Deeply-learned <i>Clustering</i>	<i>K-NEAREST NEIGHBOR</i> K-mean	(Wang et al. 2020)	Hasil : Akurasi <i>K-Nearest Neighbor</i> 69%, <i>K-Mean</i> 70%
15	Recommendation system development for fashion retail e-commerce	K-RecSys	(Hwangbo et al. 2018)	Hasil Akurasi 81 %

Mengacu pada Tabel 1. Daftar jurnal, peneliti memilih metode *K-Nearest Neighbor* dengan nilai akurasi sebesar 88,89% dan *C4.5* dengan nilai akurasi sebesar 91.33% sebagai metode yang digunakan dalam penelitian ini karena nilai akurasi pada metode tersebut memiliki nilai yang tertinggi.

Consumer online rating and review merupakan faktor yang mempengaruhi terhadap minat beli konsumen [1]. Sedangkan berdasarkan 54 responden menyatakan yang mempengaruhi minat pembelian adalah harga [18], jumlah penjualan, diskon serta deskripsi barang menggunakan algoritma CHAID [9]. Harga, kota penjual, jenis produk, jumlah penonton, dan peringkat adalah variabel independen.

SIMPULAN DAN SARAN

Dari penelitian tersebut dapat disimpulkan bahwa:

1. Faktor yang mempengaruhi kompleksitas *marketplace* yang ada di Indonesia adalah harga, jumlah penjualan, diskon, *rating and review*.
2. Metode yang paling optimal digunakan untuk analisis potensi pasar *marketplace* terhadap penjualan adalah *K-Nearest Neighbor* dan *C4.5*

DAFTAR PUSTAKA

- [1] Iprice, "Peta E-Commerce Indonesia - Q4 2021," [Accessed Maret 2022]. [Online]. Available:., 2022. <https://iprice.co.id/insights/mapofecommerce/>
- [2] N. Pransiska, A. H. Mirza, Andri., and -, "Penerapan Data Mining Prediksi Penjualan Barang Elektronik Terlaris Menggunakan Algoritma Naïve Bayes (Study Kasus : Planet Cash And Credit Cabang Muara Enim)," *Bina Darma Conf. Comput. Sci. e-ISSN 2685-2683p-ISSN 2685-2675*, pp. 2157–2169, 2017.
- [3] M. A. Muzani, M. I. A. Sukri, S. N. Fauziah, A. Fatkhurohman, and D. Ariatmanto, "Data Mining Untuk Klasifikasi Produk Menggunakan Algoritma K-Nearest Neighbor Pada Toko Online," *Pros. SISFOTEK ISSN 2597-3584*, vol. 5, no. 1, pp. 141–145, 2021, [Online]. Available: <http://seminar.iaii.or.id/index.php/SISFOTEK/article/view/273>
- [4] N. Artisya and S. P. Saragih, "Prediction Tingkat Potensi Pembeli Pada Hari Special Di Toko Mitra Shopee," *J. COMASIE Predict. E-ISSN 2715-6265*, vol. 06, no. 04, pp. 99–108, 2022, [Online]. Available: <http://ejournal.upbatam.ac.id/index.php/comasiejournal>
- [5] A. Astri, A. Fajriyati;, I. Putri;, Roghibah;, and Wa Ode Zuhayeni Madjida;, "Penerapan Algoritma K-Modes Clustering Dengan Validasi Davies Bouldin Index Pada Pengelompokkan Tingkat Minat Belanja Online Di Provinsi Daerah Istimewa Yogyakarta," *J. MSA (Mat. dan Stat. serta Apl.)*, vol. 9, no. 1, p. 24, 2021, doi: 10.24252/msa.v9i1.18555.
- [6] Mesri Silalahi, "Analisis Clustering Menggunakan Algoritma K-Means Terhadap Penjualan Produk Pada PT Batamas Niaga Jaya," *Comput. Based Inf. Syst. J. ISSN 2337-8794 / E- ISSN 2621-5292*, vol. 02, pp. 20–35, 2018, [Online]. Available: <http://ejournal.upbatam.ac.id/index.php/cbis>
- [7] E. Yunianto and T. Taryadi, "Implementasi Fuzzy Decision Making Untuk Pemilihan

- Marketplace,” *Smart Comp Jurnalnya Orang Pint. Komput. P-ISSN 2089-676X E-ISSN 2549-0796*, vol. 11, no. 1, pp. 100–104, 2022, doi: 10.30591/smartcomp.v11i1.3253.
- [8] Karsito and W. Monika Sari, “Prediksi Potensi Penjualan Produk Delifrance Dengan Metode Naive Bayes Di Pt. Pangan Lestari,” *J. Teknol. Pelita Bangsa ISSN 2407-3903*, vol. 9, no. 1, pp. 67–78, 2018.
- [9] M. I. Islami and A. Prasetio, “Analisis Dan Prediksi Penjualan Pada Marketplace Berdasarkan Pendekatan Klasifikasi Dengan Metode Pohon Keputusan (Studi Kasus Pada Data Katalog Dan Penjualan Tas Pria & Wanita Di Tokopedia Secara Nasional),” *e-Proceeding Manag. ISSN 2355-9357*, vol. 5, no. 1, pp. 30–37, 2018.
- [10] Yahya and Mahpuz, “Penggunaan Algoritma K-Means Untuk Menganalisis Pelanggan Potensial Pada Dealer SPS Motor Honda Lombok Timur Nusa Tenggara Barat,” *Infotek J. Inform. dan Teknol. e-ISSN 2614-8773*, vol. 2, no. 2, pp. 109–118, 2019.
- [11] S. Herdyansyah, E. H. Hermaliani, L. Kurniawati, and S. Rahayu, “Analisa Metode Association Rule Menggunakan Algoritma Fp-Growth Terhadap Data Penjualan (Study Kasus Toko Berkah),” *J. KHATULISTIWA Inform. p-ISSN 2339-1928 I e-ISSN 2579-633X*, vol. VIII, no. 2, pp. 2579–633, 2020, [Online]. Available: www.nusamandiri.ac.id
- [12] B. Song, W. Yan, T. Zhang, and -, “Cross-border e-commerce commodity risk assessment using text mining and fuzzy rule-based reasoning,” *Adv. Eng. Informatics*, vol. 40, no. January, pp. 69–80, 2019, doi: 10.1016/j.aei.2019.03.002.
- [13] R. V. Karthik and S. Ganapathy, “A fuzzy recommendation system for predicting the customers interests using sentiment analysis and ontology in e-commerce,” *Appl. Soft Comput.*, vol. 108, p. 107396, 2021, doi: 10.1016/j.asoc.2021.107396.
- [14] D. K. Sharma, S. Lohana, S. Arora, A. Dixit, M. Tiwari, and T. Tiwari, “E-Commerce product comparison portal for classification of customer data based on data mining,” *Mater. Today Proc.*, vol. 51, no. xxxx, pp. 166–171, 2021, doi: 10.1016/j.matpr.2021.05.068.
- [15] K. Wang, T. Zhang, T. Xue, Y. Lu, and S. G. Na, “E-commerce personalized recommendation analysis by deeply-learned clustering,” *J. Vis. Commun. Image Represent.*, vol. 71, 2020, doi: 10.1016/j.jvcir.2019.102735.
- [16] H. Hwangbo, Y. S. Kim, K. J. Cha, and -, “Recommendation system development for fashion retail e-commerce,” *Electron. Commer. Res. Appl.*, vol. 28, pp. 94–101, 2018, doi: 10.1016/j.elerap.2018.01.012.
- [17] Ichsan M . S., Jumhur ,HH, “Pengaruh Consumer Online Rating and Review Terhadap Minat Beli Konsumen Pada Marketplace Tokopedia Di Wilayah Dki Jakarta Effect of Consumer Online Rating and Review To Buying,” *e-Proceeding Manag.*, vol. 5, no. 2, pp. 1828–1835, 2018.
- [18] N. Vilano and S. Budi, “Penerapan Kansei Engineering dalam Perbandingan Desain Aplikasi Mobile Marketplace di Indonesia,” *J. Tek. Inform. dan Sist. Inf.*, vol. 6, no. 2, pp. 354–364, 2020, doi: 10.28932/jutisi.v6i2.2705.