PENERAPAN ALGORTIMA *NAÏVE BAYES* DALAM KLASIFIKASI HARGA PONSEL

Bayu Saputra, Saeful Anwar, Edi Tohidi, Heliyanti Susana, Denni Pratama
Program Studi Teknik Informatika, STMIK IKMI Cirebon
Jl. Perjuangan No. 10 B Majasem Kota Cirebon
bayustra01@gmail.com

ABSTRAK

Di jaman sekarang digital saat ini bidang teknologi berkembang pesat terutama terkait dengan beragamnya fitur yang ada pada ponsel dan tingkat spesifikasi ponsel yang semakin dinamis. Fitur ponsel semakin terus maju berkembang seperti penyimpanan memori, resolusi layar amoled, kapasitas baterai yang tahan lama menjadi aspek penentu harga ponsel. Spesifikasi ponsel semakin meningkat banyak fitur canggih yang dimiliki ponsel, maka dalam harga pun meningkat. Harga yang merupakan proses nilai tukar sangat penting digunakan dalam peluang ekonomi dan bisnis. Harga menjadi acuan pertama dalam jual beli. Penelitian ini bertujuan untuk menentukan keakuratan hasil nilai dan kinerja model klasifikasi harga ponsel yang menggunakan algoritma *naïve bayes*. Klasifikasi merupakan proses cara menemukan sebuah atribut serupa dalam sekumpulan objek dalam data yang dapat mengklasifikasikannya ke setiap kelas berbeda pada model klasifikasi yang telah ditentukan. Salah satu algoritma teknologi klasifikasi yang mudah diimplementasikan dan memiliki kecepatan mekanisme yang cepat disebut metode algoritma *Naive Bayes*. Pemodelan *statistik* dapat menggunakan metode *Naïve Bayes* dalam proses klasifikasi data. Menghitung suatu nilai *probabilitas*, metode ini dapat melatih data uji berdasarkan peristiwa data yang telah terjadi. Penelitian ini akan memanfaatkan klasifikasi yaitu *Naïve Bayes*. Pada artikel kali ini telah melakukan penelitian dengan penerapan algoritma *Naïve Bayes* dalam mengklasifikasikan harga ponsel dapat memperoleh nilai akurasi 85.50%.

Kata kunci: *Data Mining*, Klasifikasi, *Naïve Bayes*, Harga, Spesifikasi Ponsel.

1. PENDAHULUAN

Saat ini telepon genggam adalah alat komunikasi yang sangat populer karena sebagian besar orang memakainya untuk kebutuhan setiap saat seperti bersosialisasi lewat komunikasi dan mendapatkan sumber informasi. Semakin canggih teknologi, semakin banyak fitur dan spesifikasi yang ditawarkan oleh telepon genggam. Harga ponsel dipengaruhi oleh banyak fitur, termasuk seperti prosesor, bentuk yang lebih komplek, kamera yang lebih canggih, dan ram yang lebih kuat[1]. Harga adalah nilai tukar yang sangat baik untuk peluang pemasaran dan bisnis, itu menjadi acuan pertama untuk jual beli. Berdasarkan perkiraan harga yang disesuaikan, setiap pembeli harus mempertimbangkan spesifikasi ponsel yang akan mereka beli[2].

Kecerdasan buatan (*AI*), yang sekarang menjadi bidang teknis yang sangat luas yang memberi mesin kemampuan untuk menjawab pertanyaan dengan cerdas. Teknik kecerdasan buatan (*AI*) yang meliputi metode *klasifikasi*, *regresi*, *supervised learning* dan *unsupervised learning*. Dapat menggunakan pengklasifikasi apa pun seperti *Naïve Bayes* dan banyak pengklasifikasi lainnya. Karena ini adalah masalah optimasi, banyak algoritma dalam proses pemilihan fitur tersedia untuk memilih sebuah fitur terbaik dan dapat meminimalkan ukuran data. Dalam hal ini, banyak fitur yang akan dibahas, termasuk prosesor, masa pakai daya baterai, ukuran penyimpan, dan ketahanan ponsel. Faktor tersebut sangat penting untuk dipertimbangkan ketika memperkirakan harga ponsel[1]. Menurut Kristianto basis pengetahuan yaitu pemahaman atau pemahaman tentang bidang yang diperoleh melalui pendidikan dan pengalaman adalah komponen utama dari ilmu komputer yang bertujuan untuk merancang otomatisasi perilaku *smart* dalam sistem kecerdasan komputer[2]. Klasifikasi merupakan metode pembelajaran mesin, juga dikenal sebagai pembelajaran yang diawasi. Menurut Manning dkk pembelajaran tinggi adalah proses mempelajari karakteristik setiap kategori yang ada. Metode ini membuat pengklasifikasi untuk mempelajari karakteristik setiap kategori berdasarkan dokumen pelatihan yang dimilikinya [3].

Dalam penelitian lain mengenai klasifikasi komplikasi malaria dengan menggunakan *CART* (*Classification and regression Tree*) dan *Naive Bayes* menghasilkan akurasi 81.2%, dari mengantisipasi komplikasi malaria sehingga dapat membantu pasien mendapatkan pengobatan segera[4]. Penelitian sebelumnya mengenai klasifikasi penyakit kulit menggunakan algoritma *Naïve Bayes*, berdasarkan tekstur warna berbasis android, menunjukkan bahwa algoritma ini dapat mendeteksi penyakit kulit berdasarkan warna dengan akurasi sebesar 75% dan bekerja dengan baik. Sistem ini dapat mempercepat dan akurat hasil proses klasifikasi penyakit kulit[5]. Penelitian sebelumnya mengenai tentang melakukan analisis *sentimen* dari *tweet* tentang izin mengenakan masker di luar ruangan dengan menggunakan *Naïve Bayes* dan *TextBlob*. Dengan model data pengujian 20%, mereka dapat menghasilkan nilai akurasi 85%, sedangkan dengan model data pengujian 30%, mereka dapat menghasilkan nilai akurasi 83%[6].

Penelitian ini bertujuan untuk menentukan keakuratan hasil nilai dan kinerja model klasifikasi

harga ponsel yang menggunakan algoritma *naïve bayes*. Dengan menggunakan teknik ini untuk memperkirakan spesifikasi harga ponsel, peneliti diharapkan dapat menentukan keakuratan nilai klasifikasi harga berdasarkan spesifikasi ponsel[7].

Machine Learning juga dikenal sebagai pembelajaran mesin yang menyediakan teknik kecerdasan buatan seperti klasifikasi, *clustering*, *regresi*, dan sebagainya. metode *naïve bayes* terhadap penelitian ini digunakan dapat membuat model klasifikasi harga ponsel. Metode ini menghitung *probabilitas* suatu hasil. Untuk mengukur variabel *prediktor*, data dapat didistribusikan dalam berbagai bentuk, seperti gambar, *biner* 0, atau bahkan audio (*amplitudo*, *frekuensi*, *periode*). konsisten dengan hasil prediksi sebenarnya karena tidak ada skala yang sama untuk setiap variabel *independen.* Perbedaan skala ini sangat memengaruhi hasil prediksi saat melatih model pembelajaran mesin. Metode yang dapat digunakan selama proses klasifikasi pembelajaran ialah *Random Forest*, *Logistic Regression*, *Neural Networks* dan *Gradient boosting*. Penelitian ini menggunakan algoritma *naïve bayes* [3]. Pengelompokan secara berurutan item tertentu yang memiliki konsep yang mirip disebut klasifikasi. Metode *Naive Bayes* sangat efektif dalam pembelajaran terawasi karena menggabungkan dari *Teorema Bayes* dengan metode "*Naïve*" yang berarti setiap atribut atau variabelnya bersifat terbuka (*independent*) [8].

Hasil penelitian ini menunjukkan nilai pasti untuk analisis klasifikasi harga ponsel menggunakan dengan algortima *naïve bayes* dapat membagikan suatu angka hasil ketepatan yang masih belum diketahui dikenal sebagai nilai pasti. Peneliti akan memakai algoritma *naïve bayes* untuk memprediksikan klasifikasi harga sebuah ponsel. Proses ini melibatkan suatu analisis *teorema bayes* yang dikenal sebagai dari *precision*, *probability*, dan *statistics*. Penelitian akan berfokus pada sebuah analisis yang menggunakan algoritma *naïve bayes* untuk mengklasifikasikan harga berdasarkan spesifikasi ponsel. Peneliti ini dalam penerapan algortima *naïve bayes* menggunakan alat pengerjaan yang dapat mempresentasikan hasil algoritma *naïve bayes* dengan perangkat lunak yaitu *Rapid minner*[9]. Hasil yang diharapkan akan memungkinkan penerapan algoritma tersebut dan memperoleh nilai dari kinerja model *naïve bayes* (Sihombing dkk., 2021).

2. TINJAUAN PUSTAKA

2.1. *Data Mining*

Sebuah tahapan yang dilakukan pertama kali merupakan *Data mining*, yaitu pencarian informasi yang benar-benar baru. Biasanya digunakan untuk mengekstrak *esensi* pengetahuan dari sekumpulan data sehingga manusia dapat memahaminya dengan mudah. Dianalisis, divisualisasikan, dan diperbarui secara *online* dari data yang sudah ada, *pra-pemrosesan data*, model dan pertimbangan *inferensi*, *metrik* minat, dan *pasca*-pemrosesan struktur data [8].

2.2. *Knowledge discovery in database (KDD)*

Proses mengidentifikasi penjelasan yang berguna dapat menemukan entitas bentuk motif dalam data yang terdapat dalam database besar yang sebelumnya tidak diketahui *Knowledge Discovey in Database (KDD)* [8]

2.3. Klasifikasi

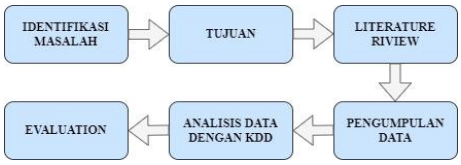
Menemukan karakteristik yang serupa dalam sekumpulan objek data dan kemudian mengklasifikasikannya ke dalam level yang tak sama berdasarkan klasifikasi setelah ditentukan merupakan proses metode klasifikasi. Target dari klasifikasi ini adalah untuk mencari model dari set pelatihan yang membedakan kelas yang sesuai, dan kemudian menggunakan model ini untuk mengklasifikasikan karakteristik kelas yang tidak diketahui. *Data Mining* menggunakan berbagai teknik klasifikasi. Untuk klasifikasi, variabel target dimasukkan ke dalam kategori [8].

2.4. *Naïve Bayes*

Algoritma *Naive Bayes* dapat melakukan sebuah proses klasifikasi data yang menggunakan pemodelan *statistik* dan salah satu metode algoritma teknologi klasifikasi yang paling mudah digunakan dan memiliki kecepatan dalam pemrosesan yang cepat. Metode ini dapat menghitung nilai hasil *probability* data uji dari peristiwa data. [11]. *Naive Bayes* merupakan teknik pembelajaran mesin yang menggunakan perhitungan *probabilitas*. Ide dasar yang digunakannya adalah *teorema bayes*, yang berarti menghitung nilai *probabilitas* untuk melakukan klasifikasi. [12]. *Teorema bayes* adalah teknik *probabilitas* dan *statistik* yang memperkirakan *probabilitas* masa depan berdasarkan pengalaman sebelumnya[7].

3. METODE PENELITIAN

Metode yang digunakan untuk memecahkan masalah terhadap penelitian ini merupakan untuk menunjukkan bagaimana model data mining dengan algoritma *naïve bayes* dapat digunakan untuk membantu mengidentifikasi klasifikasi harga ponsel. Tahapan metode penelitian dimulai dengan identifikasi masalah, tujuan, *literatur riview*, pengumpulan data, *analisis* data dengan *data mining*, dan akhirnya evaluasi.



Gambar 1. Alur Penelitian

a. Identifikasi masalah
   Proses ini mencakup identifikasi masalah untuk menentukan topik dan judul penelitian. Pada tahap ini, deskripsi aktivitas sangat penting karena membantu menentukan masalah yang akan diteliti, sehingga penelitian yang relevan dapat dilakukan tentang masalah yang relevan.
b. Tujuan
   Aktivitas yang dilakukan selama proses ini untuk menentukan tujuan penelitian yang akan dilakukan. Pada tahap ini, deskripsi aktivitas dibuat untuk menentukan tujuan penelitian yang berkaitan dengan memecahkan masalah penelitian.
c. *Literature riview*
   Dalam proses ini, tindakan yang dilakukan adalah mencari referensi yang relevan dengan penelitian yang akan dilakukan. Pada tahap ini, deskripsi aktivitas mencari referensi algoritma yang akan digunakan dalam penelitian untuk mengolah penelitian yang relevan.
d. Pengumpulan Data
   Proses ini melibatkan mencari dan mengunduh *dataset* yang relevan dan berkualitas tinggi dari situs *Kaggle*, dan menggunakan aplikasi rapid miner untuk menjalankan data. Pada tahapan ini, deskripsi aktivitas merupakan langkah terpenting dalam proses penelitian. Pertama, mencari *dataset* yang relevan berdasarkan judul. Kemudian mengunduh *dataset* sebagai bahan uji penelitian, dan kemudian menjalankan data menggunakan aplikasi *Rapid Miner*.
e. Analisis Data dengan *KDD*
   Proses ini melibatkan menganalisis data data mining untuk menentukan algoritma yang sesuai dengan penelitian. Tahap ini mencakup proses memilih algoritma pengolahan data dan langkah-langkah yang akan digunakan untuk penelitian. Analisis data mining akan menggunakan *knowledge discovery in database (KDD*). Akan menggunakan algoritma *naïve bayes* untuk algoritmanya.
f. *Evaluation*
   Mengevaluasi hasil kinerja algoritma dan model-model yang digunakan dalam penelitian ini adalah aktivitas yang dilakukan dalam proses ini. Dalam tahapan ini, deskripsi aktivitas mengukur hasil atau mengetahui hasil dari algoritma *Naive Bayes* terhadap performa model *confussion matrik* seperti *acuraccy*, *precision*, dan *recall*.

3.1. Sumber Data

Untuk penelitian ini, peneliti menggunakan data dari situs *Kaggle* berikut: https://www.kaggle.com/*datasets*/parisanahmadi/mobilepriceclassification, yang dikirim dalam format *CSV*. Data ini berisi 2001 record untuk spesifikasi ponsel dan memiliki 21 atribut. Data ini mencakup informasi tentang spesifikasi ponsel seperti *memori internal*, daya baterai, *bluetooth*, ram, *wifi*, *dual sim*, dan sebagainya.

3.2. Pengumpulan Data

Penelitian perlu memilih teknik pengumpulan data yang tepat bersama metode yang tepat juga agar penelitian dapat mengolah data dan hasil yang sesuai selama proses adaptasi. Untuk pengumpulan data dalam penelitian ini menggunakan beberapa analisis sebagai berikut:

3.2.1. Studi Pustaka

Proses yang dilakukan beserta membaca dari jurnal dan karya ilmiah, buku referensi mengenai pembelajaran yang dilakukan, dan dokumen terbitan yang sudah tersedia di perpustakaan merupakan *Studi pustaka*. Penelitian ini digunakan untuk mencari informasi sumber bacaan yang dapat menyampaikan sebuah penjelasan mengenai masalah yang diteliti.

3.2.2. *Studi literature*

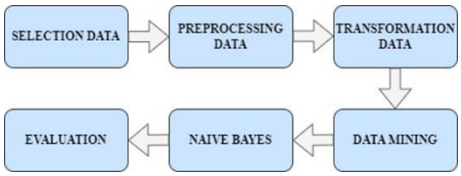
Studi literature adalah metode pengumpulan data yang mencakup melakukan penelitian, mengevaluasi data perpustakaan, membaca dan mencatat, mengolah bahan penelitian, dan sebagainya. Peneliti menggunakan metode penelitian ini untuk mengumpulkan *literature* tentang subjek dan tujuan penelitian. Teknik ini digunakan untuk memberikan penjelasan tentang berbagai teori yang terkait dengan masalah penelitian dan sebagai referensi untuk diskusi hasil penelitian

3.2.3. Pengumpulan data Secara historis

Pada *flatform* situs *website kaggle*, pengambilan data secara historis dapat dilakukan melalui berbagai kumpulan *dataset* yang mencakup informasi tentang terkait dengan judul penelitian. Peneliti dapat menjelajah dan mengunduh *dataset historis* dari sumber yang relevan dan terpercaya. Peneliti ini mengambil *dataset* dari *website* https://www.kaggle.com/*datasets*/parisanahmadi/mobilepriceclassification. Yang berbentuk format *csv* dengan nama *mobile price clasification* yang didalam nya berisi sekumpulan data spesifikasi ponsel.

3.3. Analisis Data

Analisis data merupakan serangkaian prosedur pengorganisasian, *interpretasi*, serta penyajian data untuk menarik kesimpulan dan informasi. Ini merupakan komponen penting dari metodologi penelitian dan membantu peneliti dan analis mendapatkan pemahaman yang lebih baik tentang kumpulan data dan Analisis data.



Gambar 2. Analisis data

3.3.1. Seleksi data

Pada tahap ini, data akan dipilih dan diambil untuk menentukan atribut mana yang akan digunakan dalam proses data mining. [13]

3.3.2. *Pre-processing* data

Proses pembersihan data yang dikenal sebagai *preprocessing* bertujuan untuk menghilangkan data yang tidak lengkap, salah input, tidak relevan, duplikat, dan tidak konsisten. [12]. Untuk menghapus kumpulan data yang tidak memiliki nilai (hilang), *operator replace missing* akan membersihkan nilai data yang hilang. [13]

3.3.3. Transformasi data

Data *transformation* adalah proses mengubah bentuk data yang tidak memiliki entitas menjadi bentuk data yang valid atau siap untuk proses data mining. [12]

3.3.4. *Data mining*

Pada tahap ini, data akan diproses menggunakan algoritma *naïve bayes*. *Tools rapid miner* versi 10.2.0 akan digunakan untuk memproses *data mining* terhadap klasifikasi harga ponsel. [13]

3.3.5. *Evaluation*

Pada tahap ini, tahap evaluasi adalah proses menghasilkan hasil yang dapat ditafsirkan dari proses pemodelan informasi data mining. [13]

Hasil penelitian menunjukkan bahwa algoritma *naïve bayes* berhasil memprediksi atau mengklasifikasikan harga ponsel berdasarkan data yang digunakan dalam penelitian. Berikut penjelasan untuk hasil evaluasi model menggunakan algoritma *naïve bayes*.

3.3.6. *Confusion Matrik*

Matriks konfusi bertujuan untuk memberikan informasi tentang jumlah ponsel yang diklasifikasikan dengan benar ke dalam kategori harga seperti rendah, sedang, dan tinggi, sehingga dapat memahami apakah model cenderung melakukan kesalahan tertentu, seperti mengklasifikasikan ponsel mahal sebagai ponsel murah. Misalnya, angka 0 menunjukkan harga murah, angka 1 menunjukkan harga standar, angka 2 menunjukkan harga mahal, dan angka 3 menunjukkan harga sangat mahal.[2]. Untuk menguji algoritma *Naive Bayes*, metode *confusion matrix* digunakan, yang didasarkan pada pengukuran *accuracy*, *precission*, dan *recall*, seperti yang ditunjukkan dalam persamaan 1, 2, 3.[14]

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

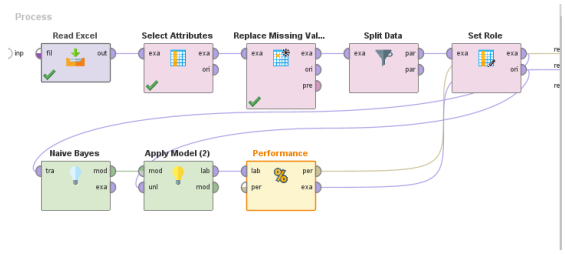
$$Recall = \frac{TN}{TP+FN} \quad (3)$$

Dengan TP adalah *True Positive*, TN adalah *True Negative*, FP adalah *False Positive*, dan FN *False Positive*. [14]

4. HASIL DAN PEMBAHASAN

4.1. Hasil

Berdasarkan teknik analisis data yang termasuk dalam metode penelitian, hasil dari klasifikasi harga ponsel menggunakan algoritma *naïve bayes* dapat dilakukan dengan menggunakan aplikasi *rapid minner*.



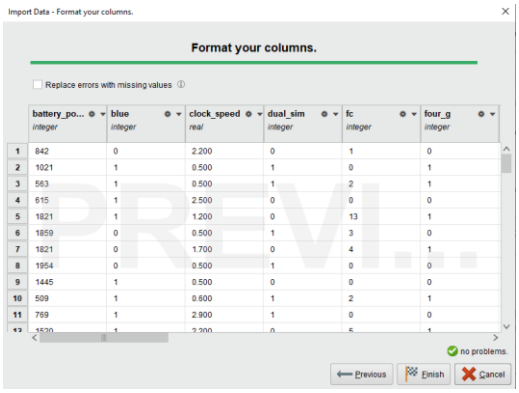
Gambar 3. Klasifikasi harga ponsel

Berdasarkan gambar 3, Proses klasifikasi menggunakan beberapa operator yang ada di *Rapid minner*, masing-masing dengan tujuan dan fungsi yang berbeda-beda dan unik. Proses klasifikasi melewati beberapa tahapan, yang terdiri dari:

4.1.1. Seleksi Data

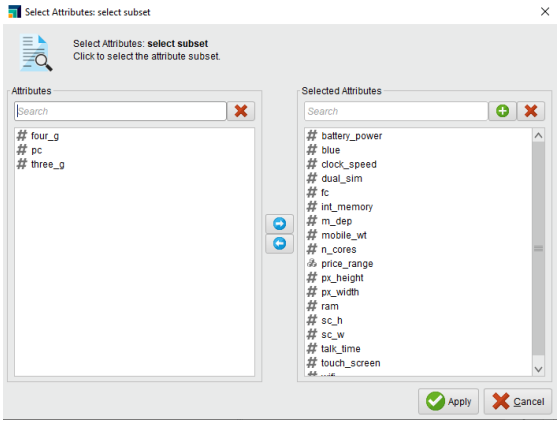
Proses pertama dalam memilih data adalah memasukkan operator "*Read Excel*", yang dapat membaca data dari *file xlsx* klasifikasi harga ponsel. Tahap berikutnya adalah mengimpor data. Pada tahap impor data, sistem akan menampilkan *preview* mengenai data sebelum memasukanya.

Tujuannya adalah untuk memeriksa kembali apakah data yang dimasukkan sudah benar dan sesuai, dan kemudian klik finish untuk menyelesaikan proses.



Gambar 4. Hasil import data

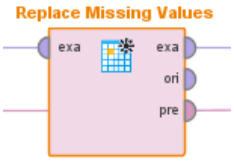
Proses selanjutnya memasukan operator *select atribut* yang berfungsi untuk memilih atribut yang akan dipakai kedalam proses data mining. Dengan atribut filter yaitu *subset*, atribut yang akan digunakan ialah 18 atribut dari data tersebut.



Gambar 5. *Parameter select atribut*

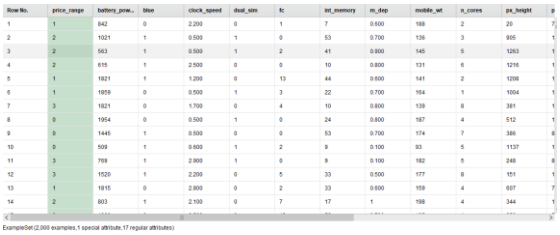
4.1.2. *Pre-processing Data*

Tidak diragukan lagi, proses pembersihan diperlukan jika ada kemungkinan adanya data yang tidak relevan. Karena data yang tidak relevan akan mengurangi akurasi pemrosesan *data mining*, dengan menggunakan *operator Replace Missing Values.*



Gambar 6. Operator *Replace Missing Values*

Dengan operator *Replace Missing Values akan* diketahui data mana yang missing atau ganda untuk diperbaiki. Setelah proses pembersihan data menggunakan operator *Replace Missing Values*, data yang digunakan adalah data lengkap, yang berarti relevan karena tidak ada data yang hilang. Dengan demikian, proses transformasi data dapat dimulai.



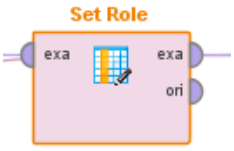
Gambar 7. Hasil *Replace Missing Values*

4.1.3. Transformasi Data

Dalam proses transformasi ini, jenis data yang dipilih diubah menjadi format yang dapat diproses dalam data mining. Dalam proses klasifikasi data mining, atribut khusus yang disebut *label* digunakan sebagai atribut penentu nilai akhir. penelitian ini akan melakukan dua transformasi data, yaitu :

a. Proses pengubahan bentuk data
   Untuk proses klasifikasi, dalam penggunaan algoritma *Naive Bayes*, bentuk data diubah dari *integer* menjadi p*olynomial*. Atribut khusus yang akan digunakan sebagai label ialah *price_range*, mengubah bentuk data menjadi *polynomial.*

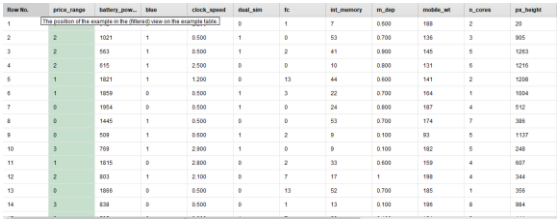
b. Proses Transformasi data yang kedua yaitu *Set role* Dalam penelitian ini, type atribut reguler diubah menjadi atribut khusus yaitu diubah menjadi label. Ini dilakukan dengan menggunakan operator *Set role* yang dapat mengubah *entitas type* data.



Gambar 8. Operator *Set rol*e

Gambar 9. *Parameter Set rol*e

Pada gambar 10, dapat dilihat perubahan atribut yang dipilih yaitu *price range* akan diubah menjadi suatu label atau atribut spesial.



Gambar 10. Hasil *Set rol*e

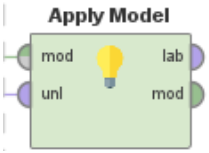
4.1.4. *Data Mining*

Aplikasi *Rapidminner* dapat digunakan untuk merancang dan mengolahan langkah-langkah proses klasifikasi yang menggunakan algoritma *naïve bayes*. Penelitian ini akan menggunakan beberapa operator yang terkait dengan penggunaan algoritma ini dalam aplikasi ini, yaitu:

a. *Naïve Bayes*
   Pada penelitian klasifikasi harga ponsel ini, algoritma *naïve bayes* digunakan. Setelah *dataset* diolah pada tahap sebelumnya, *dataset* selanjutnya akan diproses dengan operator *naïve bayes* yang tersedia pada aplikasi *Rapid Miner*.

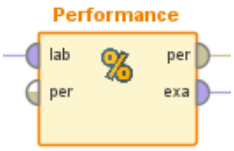
Gambar 11. Operator *naïve bayes*

b. *Apply Model*
   Proses selanjutnya memasukan operator *apply model*. Operator ini berfungsi untuk menerapkan dan memaksimalkan kinerja model dalam menjalankan pengujian *dataset*. Operator ini menghasilkan prediksi *label* pada data yang telah diproses.

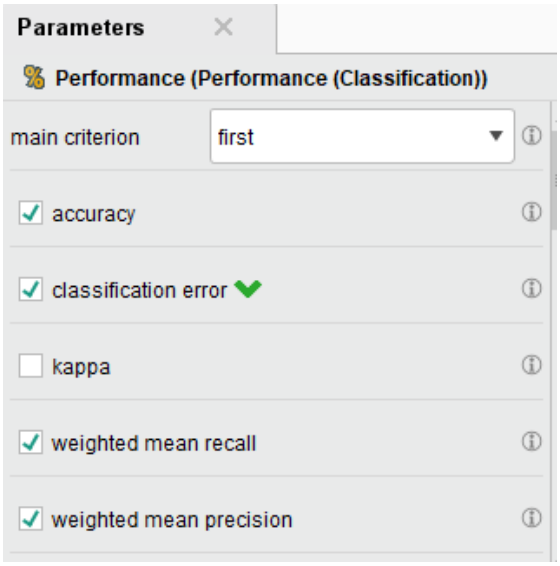


Gambar 12. Operator *apply model*

c. *Performance* (*classification*)
Selama proses penilaian penelitian ini, operator *Performance* (*clasification*) ini akan digunakan. Pada tahap ini, *output* teknik *data mining*, yang terdiri dari pola tipikal dan model *prediktif*, dievaluasi untuk mengetahui apakah *hipotesis* saat ini terpenuhi atau tidak. Jika hasilnya menunjukkan bahwa hipotesis tidak terpenuhi, ada banyak opsi yang tersedia, seperti menggunakan hasil ini sebagai umpan balik untuk meningkatkan teknik data mining atau menerima hasil ini sebagai hasil yang mungkin berguna dan bermanfaat.



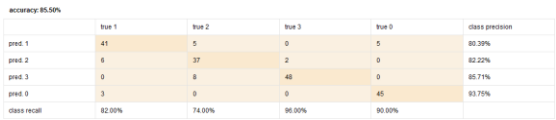
Gambar 13. Opearator *Performance*



Gambar 14. *Parameter Performance*

4.1.5. *Evaluation*

Pada tahap ini semua proses penelitian telah dilakukan semuanya sesuai dengan tahapan proses teknik analisis data. Proses evaluasi ini akan menghasilkan output nilai model dari confusion matrik yang terdiri dari nilai *accuraccy*, *recall* dan *precision*.



Gambar 15. Hasil *Performance*

Berdasarkan hasil *Performance* tentang klasifikasi harga ponsel menggunakan metode algoritma *naïve bayes* yang menghasilkan model *confusion matrik* dapat diperoleh hasil akurasi sebesar 85.50% dalam *perfomance* evaluasi model memiliki tingkat *precision* sebesar 85.52% dan *recall* sebesar 85.50%.

4.2. PEMBAHASAN

4.2.1. Penerapan Algoritma *Naïve Bayes* Dalam Mengklasifikasikan Harga Ponsel

Berdasarkan dari hasil penelitian yang telah dilakukan pada klasifikasi harga ponsel menggunakan penerapan algoritma *naïve bayes* melalui analisis statistik dan *Machine Learning*, algoritma *naïve bayes* mampu memprediksi hasil dengan akurasi yang baik. Pada metode penelitian yang digunakan untuk menerapkan algoritma *naïve bayes* menggunakan metode *knowledge discovery in database* (*KDD*). Metode ini memiliki beberapa tahapan untuk bisa menerapkan algoritma *naïve bayes* dalam mengklasifikasikan. Tahapan dalam penerapan algoritma *naïve bayes* dengan *knowledge in discovery data* (*KDD*) ialah *Data Selection* pada tahapan ini merupakan pemilihan dan pengambilan data untuk di import kedalam penelitian pengerjaan menggunakan operator *read excel*, selanjutnya *Pre-Processing data* yang merupakan persiapan data sebelum diolah dengan melakukan pemebersihan data untuk mengetahui dan menghilangkan data yang *missing* atau ganda, tahapan selanjutnya *Transformation data* pada tahapan ini akan ada pengubahan data untuk proses penerapan klasifikasi yaitu perubahan bentuk data dari *integer* ke *polynominal* dan perubahan *role* data dari *type* reguler menjadi *label*. Di tahap selanjutnya ialah *Data Mining* proses pada data mining inilah yang akan menjadi proses penerapan algoritma *naïve bayes* untuk menjalankan proses data yang telah diproses pada tahapan-tahapan sebelumnya. Proses penerapan algotima *naïve bayes* ini yang dilakukan dengan *apply model* dan *Performance (classification)* sebagai pendukung kinerja algortima *naïve bayes*.

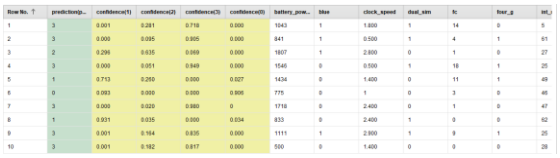
4.2.2. Evaluasi *Performance* Model Algortima *Naïve Bayes* Dalam Prediksi Harga

Hasil dari penelitian yang telah dilakukan dalam evaluasi model algoritma *naïve bayes* menghasilkan sebuah model yaitu *confusion matrik*. Model *confusion matrix* memiliki beberapa indikator penilaian untuk menghasilkan sebuah nilai yaitu *Accuracy*, *precision* dan *recall*. Dari klasifikasi harga ponsel menghasilkan nilai *accuracy* sebesar 85.50%, *precision* sebesar 85.52% dan *recall* sebesar 85.50%. Pada peneliti sebelumnya yang dilakukan oleh [4] mengenai klasifikasi malaria yang menggunakan algoritma *naïve bayes* menghasilkan nilai akurasi sebesar 81.2%. Sehingga dapat di simpulkan *naïve bayes* dalam penelitian ini mampu mengklasifikasikan dalam klasifikasi harga ponsel menggunakan

algoritma *naïve bayes* karena dapat memperoleh nilai akurasi sebesar 85.50%.

4.2.3. Rekomendasi Harga Berdasarkan Spesifikasi Ponsel

Berdasarkan hasil penelitian di atas, model algoritma *naïve bayes* yang sudah diperoleh dapat digunakan untuk merekomendasikan harga ponsel. Untuk memanfaatkan model tersebut, dapat dilakukan pengujian data baru. Rekomendasi harga yang didasarkan pada spesifikasi ponsel dapat dikategorikan ke dalam prediksi harga yang sudah ditentukan. Penelitian ini menguji coba data baru untuk menentukan spesifikasi apa yang dapat dimasukkan ke dalam prediksi harga yang sudah ditentukan.



Row No.	predictions[p...	confidence(1)	confidence(2)	confidence(3)	confidence(0)	battery_pow...	blue	clock_speed	dual_sim	fc	four_g	int_r
1	3	0.001	0.281	0.718	0.000	1043	1	1.800	1	14	0	5
2	3	0.000	0.095	0.905	0.000	841	1	0.500	1	4	1	61
3	2	0.296	0.635	0.069	0.000	1807	1	2.800	0	1	0	27
4	3	0.000	0.051	0.949	0.000	1546	0	0.500	1	18	1	25
5	1	0.713	0.260	0.000	0.027	1434	0	1.400	0	11	1	49
6	0	0.093	0.000	0.000	0.906	775	0	1	0	3	0	46
7	3	0.000	0.020	0.980	0	1716	0	2.400	0	1	0	47
8	1	0.931	0.035	0.000	0.034	833	0	2.400	1	0	0	52
9	3	0.001	0.164	0.835	0.000	1111	1	2.800	1	9	1	25
10	3	0.001	0.182	0.817	0.000	590	0	1.400	0	0	0	28

Gambar 4. 1 Prediksi rekomendasi harga

Berdasarkan dari gambar 4.14 diatas jika ingin merekomendasikan harga ponsel berdasarkan spesifikasinya dengan memanfaatkan model algoritma *naïve bayes* yang sudah tersedia maka dapat dijelaskan:

Pertama, jika ponsel yang memiliki spesifikasi dengan baterai berukuran 1043 *mAh*, mempunyai *bluetooth*, *clock_speed* berkekuatan 1.800, memiliki *dual_sim*, *front camera* sebesar 14 mp, tidak *four_g*, internal memory sebesar 5, *m_dep* memiliki 0.100 cm, *mobile_weight* berukuran 193, *n_cores* sebanyak 3, *pixel camera* sebesar 16 *mp*, *px_height* sebesar 226 cm, *px_widt* sebesar 1412 cm, *ram* sebesar 3476 mb dan *screen height* sebesar 12 dapat di kategorikan kedalam prediksi harga 3 (sangat mahal).

Kedua, jika ponsel yang memiliki spesifikasi dengan baterai berukuran 1807 *mAh*, mempunyai *bluetooth*, *clock_speed* berkekuatan 2.800, tidak *dual_sim*, *fc* sebesar 1 mp, tidak *four_g*, *internal memory* sebesar 27 gb, *m_dep* memiliki kedalaman 0.900 cm, *mobile_weight* berukuran 186, *n_cores* sebanyak 3, pc sebesar 4 *mp*, *px_height* sebesar 1270 cm, *px_width* sebesar 1366 cm, *ram* sebesar 1396 mb dan *screen height* sebesar 17 dapat di kategorikan kedalam prediksi harga 2 (mahal).

Ketiga, jika ponsel yang memiliki spesifikasi dengan baterai berukuran 1434 *mA*h, tidak mempunyai *bluetooth*, *clock_speed* berkekuatan 1.400, tidak memiliki *dual_sim*, *fc* sebesar 11, memiliki *four*_g, *internal memory* sebesar 49, m_dep memiliki kedalaman 0.500 cm, *mobile_weight* berukuran 108, *n_cores* sebanyak 6, *pc* sebesar 18 mp, *px_height* sebesar 749 cm, *px_width* sebesar 810 cm, ram sebesar 1773 mb dan *screen height* sebesar 15 dapat di kategorikan kedalam prediksi harga 1 (standar).

Keempat, jika ponsel yang memiliki spesifikasi dengan baterai berukuran 775 *mAh*, tidak mempunyai *bluetooth*, *clock_speed* berkekuatan 1, tidak memiliki *dual_sim*, fc sebesar 3, tidak *four*_g, internal memory sebesar 46, *m_dep* memiliki kedalaman 0.700 cm, *mobile_weight* berukuran 159, *n_cores* sebanyak 2, pc sebesar 3 *mp*, *px_height* sebesar 1283 cm, *px_width* 1880, ram sebesar 1495 mb dan *screen height* sebesar 7 dapat di kategorikan kedalam prediksi harga 0 (murah).

5. KESIMPULAN DAN SARAN

Dengan menggunakan algoritma *Naive Bayes* untuk mengklasifikasikan harga ponsel, hasil penelitian dapat disimpulkan Penerapan algoritma naïve bayes ini yang didukung dengan proses data mining yaitu *knowledge discovery in database* (*KDD*) yang didalam nya memiliki beberapa tahapan untuk penerapan algoritma *naïve bayes* dapat menghasilkan sistem yang cepat dan mampu mengklasifikasikan harga. Penerapan algoritma naïve bayes ini menghasilkan sebuah model yaitu *confusion matrik*. Model *confusion matrik* ini menghasilkan nilai ialah *accuracy* yang menghasilkan sebesar 85.50%, *precision* sebesar 85.52% dan *recall* sebesar 85.50%. Dengan memanfaatkan dari model algoritma *naïve bayes* yang sudah tersedia maka dapat merekomendasikan harga ponsel berdasarkan spesifikasinya. Jika ada data baru mengenai spesifikasi ponsel dan tipe ponsel dengan spesifikasi begini maka dapat mengetahui termasuk dalam prediksi harga berapa. Dengan begitu penelitian terhadap penerapan algoritma *naïve bayes* dalam klasifikasi harga ponsel memperoleh nilai akurasi sebesar 85.50%.

Dataset yang digunakan diharapkan meriset secara langsung dengan melakukan *studi kas*us untuk kevalidan data. Mencoba menggunakan penerapan metode algoritma klasifikasi lain untuk mengetahui nilai proses akurasi yang berbeda dari setiap metode algoritma lain. Mencoba menggunakan *tools* pengerjaan selain rapid minner, misalnya menggunakan *python*, *anaconda*, *orange* dan sebgainya. Agar dapat mengetahui *tools* mana yang dapat memberikan hasil nilai keakuratan data yang maksimal.

DAFTAR PUSTAKA

[1] A. Arisusanto, N. Suarna, And G. Dwilesatari, "Analisa Klasifikasi Data Harga Handphone Menggunakan Algortima Random Forest Dengan Optimize Parameter Grid," *J. Teknlogi Ilmu Komput.*, Vol. 1, No. 2, Pp. 43–47, 2023, Doi: 10.56854/Jtik.V1i2.51.

[2] Sulastria, O. D. Megawati, And A. E. Minarno, "Klasifikasi Harga Ponsel Menggunakan Support Vector Machine (Svm)," *J. Repos.*, Vol. 3, No. 4, Pp. 393–400, 2021.

[3] S. Sulastri And E. Zuliardo, "Klasifikasi Dokumen Dengan Metode Naive Bayes Terhadap Putusan Kasasi Pengadilan Tentang Merk," *J. Din. Inform.*, Vol. 12, No. 1, Pp. 45–52, 2020, Doi:

10.35315/Informatika.V12i1.8162.

[4] R. Irmanita, S. S. Prasetyowati, And Y. Sibaroni, "Classification Of Malaria Complication Using Cart (Classification And Regression Tree) And Naïve Bayes," *J. Resti (Rekayasa Sist. Dan Teknol. Informasi)*, Vol. 5, No. 1, Pp. 10–16, 2021, Doi: 10.29207/Resti.V5i1.2770.

[5] M. Furqon, Y. R. Nasution, And R. Fadillah, "Klasifikasi Penyakit Kulit Menggunakan Algoritma Naïve Bayes Berdasarkan Tekstur Warna Berbasis Android," *J. Sains Komput. Inform.*, Vol. 6, No. 1, Pp. 12–20, 2022.

[6] I. F. Ashari, F. A, M. Daffa, And S. Ali, "Sentiment Analysis Of Tweets About Allowing Outdoor Mask Wear Using Naïve Bayes And Textblob," *Indones. J. Comput. Sci.*, Vol. 12, No. 3, Pp. 1092–1103, 2023, Doi: 10.33022/Ijcs.V12i3.3238.

[7] A. Karim, F. Nurhadi, I. K. O. Setiawan, I. A. Rizky, And R. B. Manurung, "Pengaruh Normalisasi Data Pada Klasifikasi Harga Ponsel Berdasarkan Spesifikasi Menggunakan Klasifikasi Naive Bayes Dan Multinomial Logistic Regression," *J. Rekayasa Elektro Sriwij.*, Vol. 3, No. 1, Pp. 8–16, 2022.

[8] A. Firdaus And M. Walid, "Klasifikasi Kasus Covid-19 Menggunakan Model Naive Bayes Classifier," *J. Mhs. Tek. Inform.*, Vol. 6, No. 2, Pp. 583–588, 2022.

[9] D. R. Amalia, R. Narasati, And A. Faqih, "Perbandingan Hasil Klasifikasi Rasa Minuman Thai Tea Yang Paling Digemari Menggunakan K-Means Dan K-Medoids," *Pros. Semin. Nas. Unimus*, Vol. 2, Pp. 401–407, 2019.

[10] L. O. Sihombing, Hannie, And B. A. Dermawan, "Sentimen Analisis Customer Review Produk Shopee Indonesia Menggunakan Algoritma Naive Bayes Classifier," *J. Pendidik. Inform.*, Vol. 5, No. 2, Pp. 233–242, 2021, Doi: 10.29408/Edumatic.V5i2.4089.

[11] A. Rahman, "Klasifikasi Performa Akademik Siswa Menggunakan Metode Decision Tree Dan Naive Bayes," *J. Sains, Teknol. Komput. Dan Manajemen.*, Vol. 9, Pp. 22–31, 2023.

[12] H. Susana And N. Suarna, "Penerapan Model Klasifikasi Metode Naive Bayes," *J. Sist. Inf. Dan Teknol. Inf.*, Vol. 4, No. 1, Pp. 2–9, 2022.

[13] A. Nurjanah And A. Rifai, "Penerapan Algoritma K-Nearest Neighbor Untuk Klasifikasi Kelayakan Status Penduduk Miskin Di Desa Susukan Tonggoh," *J. Wahana Inform.*, Vol. 2, No. 1, Pp. 164–176, 2023.

[14] N. L. P. Merawati, A. Zulli, And Ismarmiaty, "Analisis Sentimen Dan Pemodelan Topik Pariwisata Lombok Menggunakan Algoritma Naive Bayes Dan Latent Dirichlet Allocation," *J. Resti (Rekayasa Sist. Dan Teknol. Informasi)*, Vol. 5, No. 1, Pp. 123–131, 2021, Doi: 10.29207/Resti.V5i1.2587.