



ANALISIS *TEXT MINING* KLASIFIKASI KEGIATAN KELUARGA MENGGUNAKAN ORANGE DENGAN METODE NAÏVE BAYES

Arsya Fathiarahma¹, Apriade Voutama², Taufik Ridwan³, Nono Heryana⁴

^{1,2,3,4} Sistem Informasi, Universitas Singaperbangsa Karawang
Jawa Barat, Karawang, Indonesia 41361

arsya.fathiarahma19005@student.unsika.ac.id, apriade.voutama@staff.unsika.ac.id,
taufik.ridwan@cs.unsika.ac.id, nono@unsika.ac.id

Abstract

The dual roles performed by women often lead to conflict in the family. Disputes can occur due to the lack of family roles in helping balance the work at home and office experienced by career women and mothers. The Naïve Bayes algorithm is used in this study to determine the results of applying the Naïve Bayes algorithm classification based on role, activity, age, and interest in a family. A total of 287 data records generated from the survey were used in this study and a data split of 80:20 for training and testing data using orange data mining tools. The results show that the calculation accuracy with the Naïve Bayes algorithm reaches 93%, with the conclusion that the use of orange data mining tools results in good accuracy.

Keywords: Conflict, Double role, Naïve bayes, Orange data mining tools, Woman

Abstrak

Peran ganda yang dilakukan oleh wanita kerap kali menimbulkan konflik dalam keluarga. Konflik dapat terjadi karena kurangnya peran keluarga dalam membantu keseimbangan antara pekerjaan di rumah dan kantor yang dialami oleh para wanita karir sekaligus ibu. Algoritma *Naïve Bayes* dipergunakan pada penelitian ini dengan tujuan untuk mengetahui hasil penerapan klasifikasi algoritma *Naïve Bayes* berdasarkan peran, aktivitas, umur dan minat dalam sebuah keluarga. Sejumlah 287 *data record* yang dihasilkan dari survei dipergunakan dalam penelitian ini dan dilakukan *split data* 80:20 untuk *data training* dan *data testing* dengan menggunakan *tools orange data mining*. Hasil penelitian menampilkan bahwa akurasi perhitungan dengan algoritma *Naïve Bayes* mencapai 93% dengan kesimpulan bahwa penggunaan *tools orange data mining* menghasilkan akurasi yang baik.

Kata kunci: Konflik, *Naïve bayes*, Peran ganda, *Tools orange data mining*, Wanita

1. PENDAHULUAN

Informasi yang kompleks dan variatif beberapa tahun belakangan ini dihasilkan dari perkembangan sistem informasi[1]. Berbagai peningkatan utilitas sistem informasi disebabkan karena adanya berbagai metode yang ditanamkan pada sebuah *engine*. Salah satu pemanfaatan sistem informasi yaitu ditemukannya informasi baru dari sebuah data dengan digunakannya *text mining*[2]. Terdapat beberapa penggunaan *text mining*, salah satunya yaitu *classifier* dengan *naïve bayes*[3].

Objek pada penelitian ini yaitu berawal dari perubahan pandangan masyarakat akibat tingginya gaya hidup, status sosial, ataupun adanya pergeseran budaya membuat wanita mempunyai peran ganda sebagai ibu rumah tangga sekaligus wanita karir[4]. Hal tersebut dikuatkan oleh jumlah persentase jumlah pekerja perempuan berdasarkan

status perkawinan, jenis kelamin, dan daerah tempat tinggal yang dipaparkan oleh BPS RI tahun 2018[5].

Tabel 1. Persentase Penduduk Usia 15 tahun ke Atas yang Bekerja

	Status Perkawinan				Total
	Belum Kawin	Kawin	Cerai Hidup	Cerai Mati	
Perkotaan					
Perempuan	19,96	66,51	4,39	9,14	100,00
Laki-Laki	23,04	73,33	1,78	1,85	100,00
Perdesaan					
Perempuan	10,17	76,04	3,31	10,48	100,00
Laki-Laki	19,77	76,35	1,65	2,23	100,00
Perkotaan +Perdesaan					

	Status Perkawinan				Total
	Belum Kawin	Kawin	Cerai Hidup	Cerai Mati	
Perempuan	15,39	70,96	3,89	9,76	100,00
Laki-Laki	21,52	74,73	1,72	1,03	100,00

Peran ganda yang dilakukan, akan mempengaruhi kinerja di kantor maupun di rumah yang dapat menimbulkan tekanan pada wanita yang menyebabkan terjadinya konflik atas ketidakseimbangan antara pekerjaan dan keluarga[6]. Permasalahan tersebut juga dapat disebabkan adanya stereotip yang dibuat oleh masyarakat terkait pekerjaan. Stereotip adalah menggeneralisasikan individu ke suatu kelompok tertentu dikarenakan sebuah persepsi atau sebuah penilaian[7].

Kualitas *Feminine* atau maskulin harus dimiliki oleh seseorang agar dianggap sukses dalam melakukan pekerjaan. Karena kualitas *feminine* merupakan kualitas yang diasosiasikan dengan jenis kelamin perempuan, pekerjaan rumah seperti memasak, mencuci baju, membersihkan rumah, kerap kali masyarakat menganggap bahwa perempuan lebih cocok melakukan pekerjaan tersebut[8].

Dalam penelitian ini, *dataset* yang digunakan diperoleh dari hasil survei dengan pertanyaan mengenai minat, keahlian, peran, serta aktivitas atau kegiatan keluarga oleh 80 keluarga di 4 wilayah yang berbeda. Perolehan data didapatkan sebanyak 287 yang akan dibagi menjadi 2, yaitu *data training* dan *data testing* dalam proses pemodelan. Tujuan dalam penelitian ini yaitu menampilkan hasil dari penerapan metode *Naive Bayes* dalam klasifikasi pada data yang didapatkan dari hasil survei dengan adanya penggunaan sebuah *tools* yaitu *orange data mining*.

Text mining merupakan salah satu bidang perluasan dari *data mining* yang melakukan proses pengidentifikasian pola, topik, atau atribut lainnya pada data yang berbentuk teks semi struktur atau tidak terstruktur dengan mengeksplor dan melakukan analisa yang menghasilkan temuan informasi baru[9]. Secara umum, *text mining* mengacu pada proses ekstrak informasi dari dokumen tidak terstruktur[10]. Pendefinisian *text mining* dapat dijelaskan dengan adanya penemuan informasi baru atau tidak diketahui sebelumnya oleh komputer dan informasi dari berbagai sumber yang berhasil diekstrak merupakan kunci dari proses *text mining*[11].

Salah satu algoritma yang sering digunakan terhadap klasifikasi *text mining* yaitu *Naive Bayes classifier*. *Teorema bayes* ini merupakan algoritma dengan penggunaan pengalaman sebelumnya untuk memprediksi peluang kedepannya[12]. *Naive Bayes classifier* mempunyai ciri utama yaitu independensi. Alasan teorema ini disebut naif adalah karena membuat asumsi kemunculan fitur tidak saling bergantung[13]. Bahkan jika terdapat ciri yang bergantung satu sama lain atau muncul atas

keberadaan ciri lainnya, semua sifat ini secara individual berkontribusi pada probabilitas[14]. Secara singkat, algoritma *naive bayes classification* adalah diprediksinya semua probabilitas untuk pengklasifikasian sekumpulan data algoritma yang sering

digunakan karena dapat memungkinkan dapat bekerja lebih dan kompleks dari keadaan nyata [15]

Rumus *naive bayes*:

$$P(Y|X) = \frac{P(X|Y)P(Y)}{P(X)}$$

Keterangan

X : Belum diketahui *class* pada data

A : Suatu *class* spesifik pada hipotesis data X

$P(Y|X)$: Probabilitas hipotesis berdasar kondisi

$P(Y)$: Probabilitas hipotesis

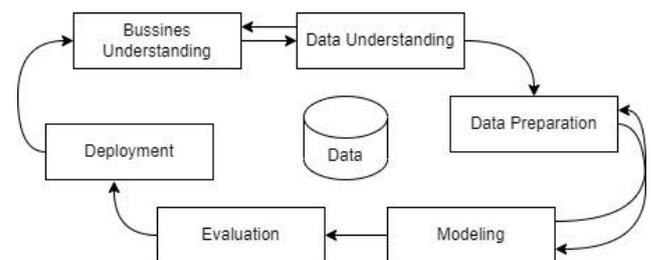
$P(X|Y)$: Probabilitas berdasarkan kondisi pada hipotesis

$P(X)$: Probabilitas Y

Berdasarkan penjelasan tujuan pada penelitian ini, *tools* yang akan digunakan untuk penerapan metode *Naive Bayes* dalam klasifikasi kegiatan keluarga yaitu *orange data mining*. Teknologi *engine open source* ini berguna dalam menganalisis dan menggambarkan data eksploratif. Di dalamnya terdapat banyak pilihan seperti pemodelan prediktif, eksperimen, maupun sistem rekomendasi sebagai bahan penelitian biomedis, genomik, dan atau pengajaran yang membuat *tools* ini mempermudah proses analisa data dengan intuitif[16].

2. METODE PENELITIAN

Kegiatan yang dilakukan sebelum penelitian adalah penyelidikan atau pencarian suatu fakta dengan teliti sesuai metode tertentu agar dapat menjelaskan dan menjawab suatu permasalahan[17]. Pada Gambar 1 merupakan tahap metode penelitian eksperimen, serta metodologi CRISP-DM yang akan dijelaskan sebagai berikut:



Gambar 1. Metode CRISP-DM

2.1. Pemahaman Bisnis (*Business Understanding*)

Pemahaman bisnis yang dimaksudkan pada penelitian ini yaitu pemahaman akan sebab dan akibat dari peran ganda dari seorang wanita yang menjadi seorang ibu, istri, sekaligus wanita karir. Seperti, adanya kesulitan untuk melakukan semua pekerjaannya, pertikaian dalam rumah tangga hanya karena seorang istri telat untuk melayani suaminya, ataupun adanya ketidakseimbangan perlakuan antar anak satu dan lainnya dalam memberikan tanggung jawab pekerjaan rumah. Maka dari itu, strategi awal yang diterapkan adalah melakukan penelitian kepada beberapa keluarga untuk mencari sekiranya apa saja pekerjaan rumah yang dapat dibagi dengan sama rata untuk mengurangi pekerjaan dari seorang wanita berperan ganda maupun pertikaian antar anak

2.2. Pemahaman Data (*Data Understanding*)

Setelah melakukan pemahaman atas permasalahan yang diangkat, selanjutnya melakukan tahapan *data understanding*. Tahapan ini dilakukan dengan menentukan atribut apa saja yang akan dipertanyakan ke keluarga-keluarga dan melakukan survei kepada mereka.

2.3. Persiapan Data (*Data Preparation*)

Tahap pada *data preparation* memiliki *preprocessing* yang akan dilakukan. *Data cleaning* merupakan salah satu *preprocessing* yang dilakukan pada *data preparation* untuk menghapus *missing value* pada salah satu atau beberapa atribut yang ada. Kegunaan dari menghapus *missing value*

pada atribut ialah agar data yang dihitung lebih akurat karena sudah tidak ada *missing value*.

2.4. Pemodelan (*Modeling*)

Setelah data telah dipersiapkan, langkah selanjutnya yaitu tahap pemodelan. Model yang digunakan pada penelitian ini yaitu *Naïve Bayes classifier*. Untuk melakukan model tersebut, dilakukannya pembagian data dengan teknik *split* antara data yang ingin dilatih (*training data*) dan data yang ingin diuji (*testing data*) sejumlah 80:20.

2.5. Evaluasi (*Evaluation*)

Tahap evaluasi yang dimaksud adalah melakukan proses tinjauan dari hasil model yang dilakukan dan perhitungannya seperti *accuracy*, *precision*, *recall*, *confusion matrix*, dan lain sebagainya

2.6. Deployment

Tahap *deployment* merupakan tahapan yang dilakukannya pembuatan laporan hasil analisis klasifikasi menggunakan model yang digunakan.

Pengumpulan data dilakukan dengan metode survei kepada beberapa keluarga. Data/atribut dalam penelitian ini diantaranya adalah Peran, Umur, Minat1, Minat2, Keahlian, Keahlian1, dan Aktivitas. Setelah data didapatkan dari survei oleh ± 80 keluarga, dihasilkan 287 *data record* yang dilakukan selama 3 hari. Hasil *data record* survei digambarkan pada Gambar 2.

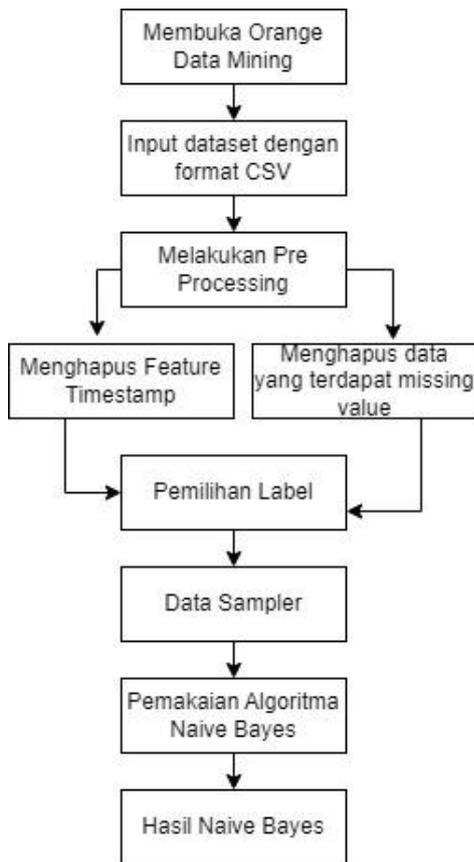
	Timestamp	Peran	Umur	Minat	Minat1	Minat2	Keahlian	Keahlian1	Aktivitas
0	21/05/2022 12:39:46	Anak	21-55	Alam	Hiburan	Animasi	Masak	Manajemen	Membersihkan Rumah
1	21/05/2022 12:44:19	Ayah	21-55	Makanan	Olahraga	Binatang	Swakarya/DIY	Berkebun	Membersihkan Rumah
2	21/05/2022 12:46:16	Anak	15<	Makanan	Animasi	Binatang	Swakarya/DIY	Sosialisasi	Belanja Kebutuhan
3	21/05/2022 12:49:24	Ibu	21-55	Makanan	Olahraga	Hiburan	Masak	Sosialisasi	Memasak
4	21/05/2022 13:03:39	Anak	15-20	Makanan	Hiburan	Keuangan	Masak	NaN	Membersihkan Rumah
...
282	25/05/2022 7:31:20	Anak	15<	Olahraga	Alam	Binatang	Berkebun	NaN	Mencuci Piring
283	25/05/2022 19:23:51	Anak	21-55	Makanan	Hiburan	Edukasi	Sosialisasi	NaN	Membersihkan Rumah
284	25/05/2022 19:28:45	Ibu	21-55	Makanan	Hiburan	Keuangan	Masak	Berkebun	Memasak
285	25/05/2022 19:29:51	Ayah	21-55	Teknologi	Hiburan	Edukasi	Otomotif	Kelistrikan	Memperbaiki Rumah/Kelistrikan/Perabotan
286	26/05/2022 17:43:45	Anak	15<	Makanan	Teknologi	Hiburan	Masak	Manajemen	Memasak

287 rows × 9 columns

Gambar 2. *Dataset* yang Dihasilkan dari Survei

Basis dari tahapan penelitian ini adalah dilakukannya metode CRISP-DM. Namun, sebelum digunakannya algoritma *naïve bayes*, perlu dilakukannya *preprocessing* seperti *Cleaning*. Proses *cleaning* pada penelitian ini digunakan untuk menghapus *feature timestamp* pada

dataset dan menghapus data yang terdapat *missing value*. Alur lengkap dari tahapan penelitian yang digunakan dijelaskan pada Gambar 3.



Gambar 3. Tahapan Pemodelan

3. HASIL DAN PEMBAHASAN

Terdapat beberapa aplikasi *text mining* yang dapat digunakan untuk perhitungan secara otomatis. Namun, penelitian ini memfokuskan perhitungan *text mining* dengan metode *Naive Bayes classifier* menggunakan aplikasi atau *tools Orange Data Mining*. Berikut hasilnya:

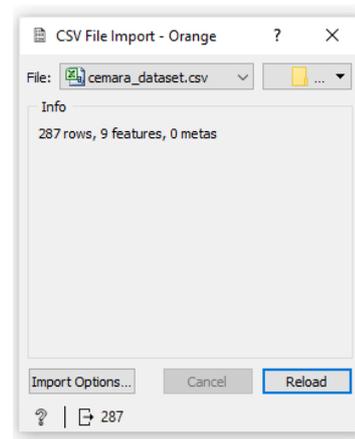
3.1. Persiapan Dataset

Hasil survei yang sudah dilakukan melalui *google form* menghasilkan 287 data yang akan dilakukan analisa. Di bawah ini terdapat beberapa tampilan dalam mempersiapkan *dataset*. Gambar 4 adalah *widget* untuk melakukan *import dataset* dengan format *csv*, Gambar 5 tampilan ketika *import file* dan Gambar 6 merupakan tampilan apabila *dataset* sudah di *import*



CSV File Import

Gambar 4. Widget Import File



Gambar 5. Import File CSV

Column type	1	2	3	4	5	6	7	8	9
1	Timestamp	Peran	Umur	Minat	Minat1	Minat2	Keshian	Keshian1	Aktivitas
2	21/05/2022 ...	Anak	21-35	Alam	Hiburan	Animasi	Masak	Manajemen	Membersihkan ...
3	21/05/2022 ...	Ayah	21-35	Makanan	Olahraga	Binatang	Swakarya/DIY	Berkebun	Membersihkan ...
4	21/05/2022 ...	Anak	15+	Makanan	Animasi	Binatang	Swakarya/DIY	Sosialisasi	Belanja ...
5	21/05/2022 ...	Ibu	21-35	Makanan	Olahraga	Hiburan	Masak	Sosialisasi	Memasak
6	21/05/2022 ...	Anak	15-20	Makanan	Hiburan	Keuangan	Masak	Kebijakan	Membersihkan ...
7	21/05/2022 ...	Ibu	21-35	Makanan	Alam	Kesehatan	Masak	Berkebun	Mengunjungi Anak
8	21/05/2022 ...	Ayah	21-35	Olahraga	Hiburan	Keuangan	Swakarya/DIY	Kebijakan	Memperbaiki ...
9	21/05/2022 ...	Anak	15+	Alam	Hiburan	Animasi	Swakarya/DIY	Sosialisasi	Mencuci Piring
10	21/05/2022 ...	Ibu	21-35	Olahraga	Hiburan	Keuangan	Masak	Berkebun	Memasak
11	21/05/2022 ...	Ayah	21-35	Teknologi	Hiburan	Animasi	Sosialisasi	Otomotif	Memperbaiki ...

Gambar 6. Tampilan Dataset di Orange

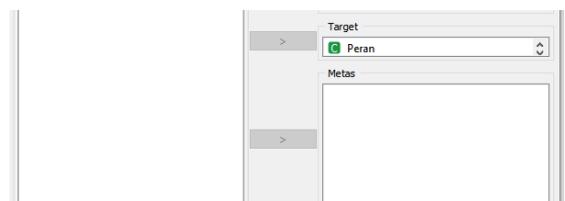
3.2. Pemilihan Label

Langkah yang dilakukan setelah *import file csv* adalah pemilihan label dan Gambar 7 merupakan tampilan dari *widget* pemilihan kolom yang akan dijadikan label. Dalam penelitian ini, peran merupakan *feature* yang dipilih untuk dijadikan label, dipaparkan melalui Gambar 8.



CSV File Import Select Columns

Gambar 7. Widget Select Columns



Gambar 8. Pemilihan Label

3.3. Preprocessing

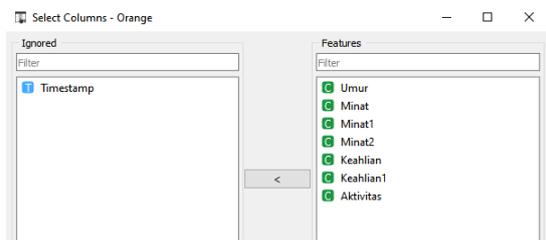
Setelah dilakukannya pemilihan *feature* untuk dijadikan label, tahap selanjutnya ialah proses *preprocessing* yaitu *cleaning*. Gambar 9 merupakan tampilan dari *widget preprocessing*.



Gambar 9. Widget Preprocessing

a) Hapus *feature timestamp*

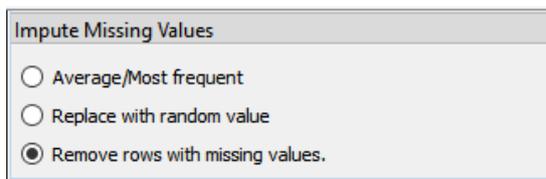
Dari total 9 *features* pada *dataset*, *feature* per sudah dipilih untuk menjadi label dan *feature timestamp* dipilih untuk dihapus dalam penelitian ini. Total *features* yang akan digunakan sekarang adalah 7 dan Gambar 10 menjelaskan mengenai apa saja yang menjadi *features* dan yang akan dihapus



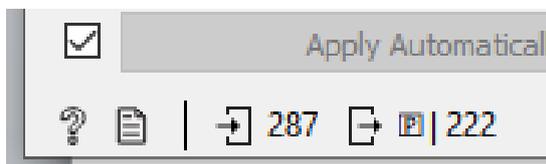
Gambar 10. Tampilan Hapus *Timestamp*

b) Hapus data yang terdapat *missing value*

Preprocessing dilanjut dengan penghapusan data/rows yang terdapat *missing value*, dijelaskan pada Gambar 11. Hasil dari penghapusan *missing value* digambarkan pada Gambar 12 yaitu terdapat 222 data yang akan di *split* untuk *training* dan *testing*



Gambar 11. Hapus *Missing Value*

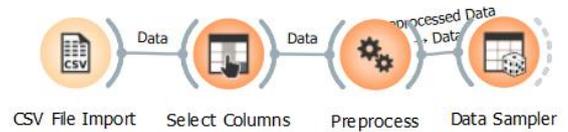


Gambar 12. Hasil Data yang Sudah Dihapus

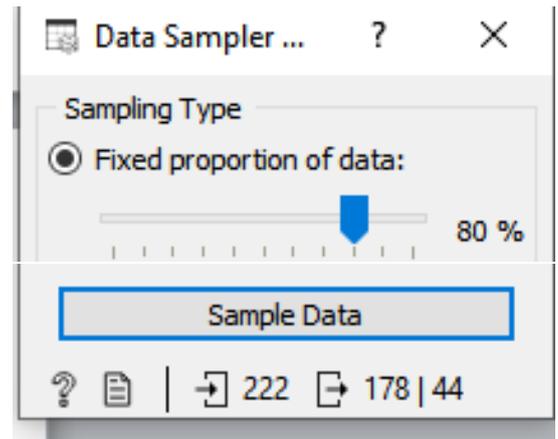
3.4. *Data Sampler*

Data sampler atau pemilihan data ini ialah proses pemilihan jumlah data yang akan dilatih (*data training*) dan jumlah data yang akan diuji (*data testing*). Seperti yang ditulis sebelumnya, proporsi dari *data training* dan *data testing* yaitu 80:20 dengan hasil 178 data yang

akan di *training* dan 44 data yang akan di *testing*. Gambar 13 merupakan tampilan *widget data sampler* dan Gambar 14 ialah tampilan untuk pengaturan proporsi dan hasil dari proporsi *data training* dan *data testing*.



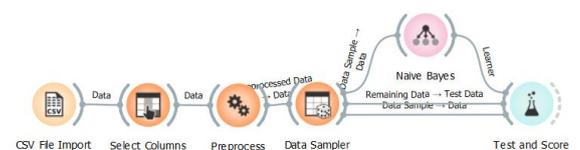
Gambar 13. *Widget Data Sampler*



Gambar 14. Proporsi *Data Train* dan *Test*

3.5. *Naïve Bayes*

Proporsi data sudah ditentukan dengan dilakukannya *split data*, maka langkah selanjutnya ialah melakukan pemodelan dengan *Naïve Bayes*. Gambar 15 merupakan tampilan secara utuh bagaimana pemodelan *Naïve Bayes* dilakukan pada *tools orange data mining*



Gambar 15. Pemodelan *Naïve Bayes*

3.6. Hasil

Evaluation results for target (None, show average over classes)					
Model	AUC	CA	F1	Precision	Recall
Naive Bayes	0.938	0.886	0.885	0.900	0.886

Evaluation results for target Anak					
Model	AUC	CA	F1	Precision	Recall
Naive Bayes	0.937	0.932	0.870	1.000	0.769

Gambar 16. Hasil Penelitian (a)

Evaluation results for target Ayah

Model	AUC	CA	F1	Precision	Recall
Naive Bayes	0.948	0.932	0.870	0.909	0.833

Evaluation results for target Ibu

Model	AUC	CA	F1	Precision	Recall
Naive Bayes	0.932	0.909	0.905	0.826	1.000

Gambar 17. Hasil Penelitian (b)

Hasil akurasi yang didapatkan dari penggunaan *tools orange data mining* menunjukkan 93% yang berarti mempunyai hasil yang baik. Selain akurasi, *precision* yang didapatkan pun mendapatkan hasil yang baik seperti yang ditampilkan pada Gambar 16 (a) dan (b).

Predictions - Orange

Show probabilities for Classes in data and model

	Naive Bayes	Peran	Umur	Minat	Minat1	Minat2	Keahlian	Keahlian1	Aktivitas
1	0.54 : 0.43 : 0.03 – Anak	Anak	21-55	Tekno...	Edukasi	Keuan...	Sosialis...	Manaje...	Members...
2	1.00 : 0.00 : 0.00 – Anak	Anak	15-20	Makanan	Tekn...	Hiburan	Masak	Sosialis...	Mencuci...
3	0.00 : 0.00 : 1.00 – Ibu	Ibu	21-55	Makanan	Olahr...	Keuan...	Masak	Berkebu...	Memasak
4	0.16 : 0.08 : 0.76 – Ibu	Anak	21-55	Hiburan	Keua...	Keseha...	Masak	Manaje...	Belanja K...
5	0.01 : 0.01 : 0.98 – Ibu	Ibu	21-55	Olahraga	Hibur...	Keseha...	Masak	Berkebu...	Memasak
6	0.03 : 0.00 : 0.97 – Ibu	Ibu	21-55	Makanan	Edukasi	Keuan...	Masak	Swakar...	Memasak
7	0.17 : 0.76 : 0.07 – Ayah	Ibu	21-55	Makanan	Hibur...	Keuan...	Sosialis...	Otomotif	Members...
8	1.00 : 0.00 : 0.00 – Anak	Anak	15-20	Tekno...	Hibur...	Animasi	Masak	Sosialis...	Mencuci...
9	0.04 : 0.84 : 0.12 – Ayah	Anak	21-55	Edukasi	Keua...	Keseha...	Sosialis...	Manaje...	Members...
10	0.90 : 0.09 : 0.01 – Anak	Anak	15-20	Makanan	Olahr...	Keseha...	Sosialis...	Otomotif	Members...

Gambar 18. Hasil Prediksi

Gambar 17 merupakan beberapa hasil dari prediksi *data testing* yang digunakan dalam penelitian. Pada tabel sebelah kiri yang bertuliskan *Naive Bayes*, merupakan tabel hasil prediksi dari *Naive Bayes* itu sendiri dan angka di data tersebut merupakan tingkat optimisme. Lalu pada sebelah kanan dari tabel *Naive Bayes* merupakan *data actual*.

Predicted

	Anak	Ayah	Ibu	Σ
Actual Anak	10	1	2	13
Actual Ayah	0	10	2	12
Actual Ibu	0	0	19	19
Σ	10	11	23	44

Gambar 19. Confussion Matrix

Gambar 18 menampilkan hasil *confussion matrix* dari 44 *data testing* yang digunakan. Klasifikasi yang bernilai valid sejumlah 10 dari 10 dari data anak, 10 dari 11 dari data ayah, dan 19 dari 23 dari data ibu[7].

4. KESIMPULAN

Berdasarkan hasil dan pembahasan yang sudah dievaluasi, penelitian ini menghasilkan model yang terbilang baik dengan hasil rata-rata akurasi (AUC) yang didapatkan sejumlah 93%, *class precision* “anak” didapatkan 100%,

class precision “ayah” didapatkan 90% dan *class precision* “ibu” didapatkan 82%.

Ucapan Terima Kasih

Puji syukur kepada Tuhan YME telah memberikan kelancaran sehingga penulis dapat menyelesaikan penelitian ini dan penulis mengucapkan terima kasih atas keterlibatan dosen maupun teman-teman dalam penelitian ini.

DAFTAR PUSTAKA

- [1] M. F. Andriansyah, D. Yusup, and A. Voutama, “Menggunakan Metode *Naive Bayes* Berbasis Website Web-Based Expert System of Covid-19 Early Detection Using *Naive Bayes* Method,” *J. Inf. Technol. Comput. Sci.*, vol. 4, no. 2, pp. 446–455, 2021.
- [2] M. T. Informatika and U. Amikom, “Analisis Pembobotan Kata pada Klasifikasi Text Mining,” vol. 3, no. 2, pp. 179–184, 2019.
- [3] A. V. Sudiantoro *et al.*, “Analisis Sentimen Twitter Menggunakan Text Mining Dengan Algoritma *Naive Bayes* Classifier,” *Din. Inform.*, vol. 10, no. 2, pp. 398–401, 2018.
- [4] D. M. Putri, M. T. Sari, and S. U. Khumaidatul, “Pengaruh Konflik Peran Ganda Terhadap Fear of Success Pada Wanita Karir yang Telah Menikah di Kota Samarinda,” *J. Psikol.*, vol. 9, no. 1, pp. 1–10, 2021.
- [5] Badan Pusat Statistik dan Kementerian Pemberdayaan Perempuan dan Perlindungan Anak, “Pr o f i l e r e m p u ON,” *Profil Peremp. Indones.*, 2019, [Online]. Available: https://www.kemenpppa.go.id/lib/uploads/list/b4bd-c-profil-perempuan-indonesial-_2019.pdf
- [6] L. Cahyadi and D. Prastyani, “Work Life Balance, Stres Kerja Dan Konflik Peran Terhadap Kepuasan Kerja Pada Wanita Pekerja,” *J. Ekon. J. Econ.*, vol. 11, no. 2, 2020, doi: 10.47007/jeko.v11i2.3576.
- [7] D. Dianti and W. Cahyati, “Persepsi Masyarakat Pada Program Studi Ilmu Komunikasi,” *Buana Komun. (Jurnal Penelit. dan Stud. Ilmu Komunikasi)*, vol. 2, no. 2, p. 116, 2022, doi: 10.32897/buanakomunikasi.2021.2.2.1134.
- [8] R. Safira, “Captain Marvel: Dominasi Maskulin Dalam Kesetaraan Gender Perempuan,” *Interak. Online*, vol. 10, no. 3, pp. 619–631, 2022, [Online]. Available: <https://fisip.undip.ac.id/>
- [9] M. Prasojo and J. Triwidianti, “Prediksi Prestasi Siswa SMK Masuk Pasar Kerja Menggunakan Teknik Data Mining (Studi Kasus SMKN 1 Kota Agung Timur Tanggamus, Lampung),” *Semin. Nas. Has. Penelit. dan Pengabd. Masy.*, pp. 134–150, 2021.
- [10] M. M. Yin, M. R. Mahmoudi, and A. Abbasalizadeh, “Analysis of mystical concepts in

- Khaghani's Divan," *Digit. Scholarsh. Humanit.*, vol. 35, no. 2, pp. 485–491, 2020, doi: 10.1093/LLC/FQZ015.
- [11] Normah, B. Rifai, S. Vambudi, and R. Maulana, "Analisa Sentimen Perkembangan Vtuber Dengan Metode Support Vector Machine Berbasis SMOTE," *J. Tek. Komput. AMIK BSI*, vol. 8, no. 2, pp. 174–180, 2022, doi: 10.31294/jtk.v4i2.
- [12] W. P. Nurmayanti, "Penerapan Naive Bayes dalam Mengklasifikasikan Masyarakat Miskin di Desa Lepak," *Geodika J. Kaji. Ilmu dan Pendidik. Geogr.*, vol. 5, no. 1, pp. 123–132, 2021, doi: 10.29408/geodika.v5i1.3430.
- [13] F. V. Sari and A. Wibowo, "Analisis Sentimen Pelanggan Toko Online Jd.Id Menggunakan Metode *Naive Bayes* Classifier Berbasis Konversi Ikon Emosi," *J. SIMETRIS*, vol. 10, no. 2, pp. 681–686, 2019.
- [14] M. Nizam Fadli, I. Sudahri Damanik, E. Irawan, S. Tunas Bangsa, and S. Utara, "KLIK: Kajian Ilmiah Informatika dan Komputer Penerapan Metode Naive Bayes Dalam Menentukan Tingkat Kenyamanan Pada Rumah Sakit Terhadap Pasien," *Media Online*, vol. 2, no. 3, pp. 117–122, 2021, [Online]. Available: <https://djournals.com/klik>
- [15] D. Wilandani and Purwanto, "Penerapan Algoritma Naive Bayes dalam Mengklasifikasikan Media Sosial untuk Mengamati Trend Kuliner," vol. 8, no. 1, pp. 31–39, 2022.
- [16] S. I. Nurhafida and F. Sembiring, "Analisis Text Clustering Masyarakat di Twiter Mengenai Mcdonald'sxpbs Menggunakan Orange Data Mining," *Semin. Nas. Sist. Inf. dan Manaj. Inform.*, vol. 1, no. 1, pp. 28–35, 2021.
- [17] A. Voutama, I. Maulana, and N. Ade, "Interactive M-Learning Design Innovation using Android-Based Adobe Flash at WFH (Work From Home)," *Sci. J. Informatics*, vol. 8, no. 1, pp. 127–136, 2021, doi: 10.15294/sji.v8i1.27880.