



Artikel

Penerapan *Data Mining* Untuk Merekomendasikan Seri Produk *NAS* Kepada Calon Konsumen Toko *Storage* Menggunakan Algoritma *Multinomial Naïve Bayes*

Fernando Verdy Sunata¹, Susanto Hariyanto², Hartana Wijaya³

^{1,2,3}, Universitas Buddhi Dharma, Teknik Informatika, Banten, Indonesia

SUBMISSION TRACK

Received: August 28, 2022

Final Revision: September 12, 2022

Available Online: September 15, 2022

KEYWORD

Data Mining, *Multinomial Naïve Bayes*,
Network Attached Storage

KORESPONDENSI

Phone: 0895411868414

E-mail: fernandoverdysunata18@gmail.com

A B S T R A K

Toko *Storage* merupakan nama dagang yang digunakan oleh PT. Distributor Trimitra Indonesia untuk menjual berbagai macam produk *NAS* (*Network Attached Storage*). Banyaknya produk *NAS* yang dijual dengan harga dan spesifikasi yang berbeda-beda, terkadang membuat bingung bahkan membuat calon konsumen kesulitan dalam memilih produk *NAS* yang tepat. Sehingga tidak jarang dari mereka yang bertanya mengenai rekomendasi *NAS* kepada *admin* toko. Proses pemberian rekomendasi dilakukan melalui sesi tanya jawab terkait dengan kebutuhan *NAS*. Proses pemberian rekomendasi terkadang memakan waktu yang lama karena harus menunggu jawaban dari calon konsumen. Karena itu, dilakukan penelitian yang bertujuan untuk merancang sebuah sistem yang mampu memberikan rekomendasi seri produk *NAS* kepada calon konsumen dengan menerapkan metode *data mining* dan algoritma *multinomial naïve bayes* (*MNB*). Hasil dari penerapan metode dan algoritma yang digunakan terbukti berhasil diimplementasikan pada data yang digunakan, hal ini dibuktikan dari hasil pengujian dan evaluasi yang dilakukan menggunakan bantuan aplikasi *Weka* yang menghasilkan nilai akurasi sebesar 95,5556%. Hasil akhir dari penelitian ini berupa rancangan sistem rekomendasi seri produk *NAS* berbasis *web* yang dapat digunakan oleh pengguna untuk mendapatkan rekomendasi seri produk *NAS* secara cepat dan tepat, hanya dengan memasukan kriteria *NAS* yang dicari.

PENGANTAR

Seiring berkembangnya perusahaan, data-data yang dimiliki oleh perusahaan juga akan terus bertambah. Hal tersebut tentunya membuat pihak perusahaan harus

menyediakan media penyimpanan data yang cukup untuk menampung semua data-data yang dimiliki dalam mendukung kegiatan operasional bisnis mereka.

Toko *Storage* merupakan salah satu toko *online* yang menjual dan menyediakan

berbagai macam perangkat *NAS* untuk memenuhi kebutuhan penyimpanan data. Banyaknya produk *NAS* yang dijual dengan spesifikasi dan harga yang berbeda-beda, terkadang membuat bingung bahkan membuat beberapa calon konsumen kesulitan dalam memilih produk *NAS* yang tepat dan sesuai untuk kebutuhan mereka. Karena hal tersebut, tidak jarang juga dari mereka yang bertanya mengenai rekomendasi produk maupun seri produk *NAS* kepada *admin* toko.

Pemberian rekomendasi produk maupun seri produk kepada calon konsumen saat ini masih dilakukan melalui sesi tanya jawab. Akan tetapi, kegiatan pemberian rekomendasi ini terkadang memakan waktu yang lama. Hal ini karena pihak *admin* harus mengetahui secara lengkap apa saja kebutuhan *NAS* yang dicari oleh calon konsumen sebelum dapat merekomendasikan produk atau seri produk *NAS*.

Oleh sebab itu, maka dilakukan penelitian yang bertujuan untuk merancang sebuah sistem yang mampu memberikan rekomendasi seri produk *NAS* dengan memanfaatkan metode *data mining* dengan menggunakan algoritma *multinomial naïve bayes (MNB)*.

I. METODE

1.1 Data Mining

Data mining merupakan sebuah proses semi otomatis yang memanfaatkan teknik statistik, matematika, kecerdasan buatan, serta pembelajaran mesin untuk mengekstraksi dan mengidentifikasi suatu informasi pengetahuan potensial yang tersembunyi di dalam *database* [1].

Output atau keluaran dari *data mining* sendiri dapat digunakan sebagai alternatif dalam pengambilan suatu keputusan maupun memperbaiki pengambilan suatu keputusan di masa yang akan datang [2].

Data mining memiliki begitu banyak algoritma atau teknik atau metode dalam menggali maupun mencari pengetahuan atau informasi. Setiap algoritma atau teknik atau metode tersebut memiliki fungsi serta tujuan yang berbeda-beda [3].

Terdapat tujuh tahapan dalam proses *data mining*, tahapan-tahapan tersebut meliputi [4]:

1. Pembersihan Data

Tahap dimana data akan diproses dengan cara dibersihkan untuk menghilangkan data yang tidak konsisten atau tidak lengkap.

2. Integrasi Data

Tahap integrasi data merupakan tahap dilakukannya penggabungan data dari beberapa sumber atau *database* ke dalam suatu *database* baru maupun file baru.

3. Seleksi Data

Tahap seleksi data merupakan tahap yang dilakukan untuk menyeleksi data atau atribut penting yang akan digunakan untuk proses *mining*.

4. Transformasi Data

Tahapan transformasi data merupakan tahap yang dilakukan untuk mentransformasi atau mengubah data menjadi bentuk atau format yang sesuai untuk di proses *mining* dengan algoritma *data mining* yang digunakan.

5. Proses Mining

Proses *mining* merupakan proses menambang atau menggali informasi tersembunyi dari data yang sudah diproses pada tahapan sebelumnya dengan menggunakan algoritma *data mining* yang digunakan.

6. Evaluasi Pola

Tahap evaluasi pola merupakan tahap yang dilakukan untuk mengevaluasi proses *data mining* yang dilakukan. Tujuan dari tahap evaluasi ini adalah untuk menilai apakah hasil hipotesa dari proses *mining* data yang dilakukan sudah tercapai atau belum.

7. Presentasi Pengetahuan

Presentasi pengetahuan merupakan tahapan yang bertujuan untuk mempresentasikan atau memvisualisasikan hasil pengetahuan yang didapat dari proses *data mining*.

1.2 Klasifikasi

Klasifikasi merupakan salah satu kelompok dalam *data mining*. Klasifikasi adalah sebuah kegiatan dalam menjalankan

suatu proses pengelompokan dengan mengacu pada karakteristik tertentu [5].

1.3 Multinomial Naïve Bayes

Multinomial naïve bayes merupakan algoritma yang mengasumsikan kemunculan kata-kata dalam sebuah dokumen adalah independen, tanpa mempertimbangkan urutan kata atau konteks informasi dari dokumen tersebut. Selain itu, algoritma ini juga memperhitungkan frekuensi kemunculan kata dalam dokumen [6].

Algoritma *multinomial naïve bayes* merupakan pengembangan dari algoritma *naïve bayes* yang sering digunakan untuk mengklasifikasikan data berupa teks [7].

Terdapat beberapa langkah yang harus dilakukan untuk mengklasifikasi suatu data menggunakan algoritma *multinomial naïve bayes*:

1. Menghitung Probabilitas Class Terhadap Dokumen

Langkah pertama yang harus dilakukan adalah menghitung probabilitas *class* terhadap dokumen dengan menggunakan rumus:

$$P(c) = \frac{N(c)}{N}$$

Keterangan:

- $P(c)$: Nilai probabilitas kelas terhadap dokumen.
- $N(c)$: Total jumlah dokumen pada kelas.
- N : Total jumlah dokumen.

2. Membuat Term-document Matrix

Berikutnya adalah membuat *term-document matrix* yang berisikan, jumlah kata pada seluruh dokumen, jumlah kata unik pada seluruh dokumen, serta jumlah kata pada setiap kelas.

3. Menghitung Probabilitas Kata Unik Pada Setiap Class

Langkah ketiga menghitung probabilitas kata unik pada setiap *class* dengan menerapkan teknik *laplace smoothing* untuk mencegah munculnya *zero probability*. Sehingga didapatkan hasil rumus sebagai berikut:

$$P(w|c) = \frac{\text{count}(w, c) + 1}{\text{count}(c) + |V|}$$

Keterangan:

- $P(w|c)$: Nilai probabilitas kata terhadap kelas.
- $\text{count}(w, c)$: Total jumlah kemunculan kata pada kelas.
- $\text{count}(c)$: Total jumlah kata pada kelas.
- $|V|$: Nilai mutlak dari jumlah kata unik pada seluruh dokumen.

4. Menghitung Probabilitas Dokumen Terhadap Class

Langkah keempat yang harus dilakukan adalah menghitung nilai probabilitas dokumen terhadap kelas dengan menggunakan persamaan rumus sebagai berikut:

$$P(c|d_{(n)}) = P(c) \times \prod P(w|c)$$

Keterangan:

- $P(c|d_{(n)})$: Nilai probabilitas kalimat terhadap kelas.
- $P(c)$: Nilai probabilitas kelas terhadap dokumen.
- $\prod P(w|c)$: Perkalian beruntun dari nilai probabilitas kata terhadap kelas.

1.4 Weka

Aplikasi *WEKA* atau *Waikato Environment for Knowledge Analysis* adalah sebuah perangkat lunak *machine learning* populer yang dibuat dengan bahasa pemrograman *Java* [8].

Aplikasi *Weka* memiliki beberapa kelebihan, diantaranya memiliki banyak algoritma *data mining* dan pembelajaran mesin, kemudahan dalam menggunakannya, algoritma yang ada selalu *up-to-date* [9].

1.5 Flask

Flask merupakan sebuah *micro-framework* yang dibuat dengan menggunakan bahasa pemrograman *python* yang tidak memiliki banyak *tools* dan *library*. *Flask* dapat dimanfaatkan untuk membantu dalam mengembangkan perangkat lunak menjadi lebih efisien [10].

1.6 Network Attached Storage (NAS)

Network Attached Storage (NAS) merupakan jenis media penyimpanan data yang terhubung ke dalam suatu jaringan dan memiliki sistem operasi sendiri. NAS memungkinkan para penggunanya untuk dapat mengakses, mengolah, dan mengedit file yang disimpan di dalamnya secara bersama-sama.

II. HASIL

Dari Data yang sudah melewati tahap *pre-processing* data kemudian dilakukan tahap klasifikasi menggunakan aplikasi *Weka*.

Tabel 1. Tabel Contoh Hasil Pre-Processing Data

No	Kriteria	Label
1	1jt sampai 20jt storage dan backup tower 2 bay tidak ada 2 gb 1 sampai 1000	home
2	20jt sampai 40jt storage dan backup tower 9 bay tidak ada 32 gb 1001 sampai 2000	smb
3	20jt sampai 40jt storage dan backup tower 8 bay ya intel tidak 16 gb lebih dari 8000	enterprise

Berikut ini adalah hasil klasifikasi dari data yang sudah melewati tahap *pre-processing* data menggunakan aplikasi *Weka*.

Tabel 2. Tabel Hasil Klasifikasi Aplikasi Weka

No	Actual	Predicted Error	Prediction
1	home	home	1
2	home	home	1
3	home	home	1
4	home	home	1
5	home	home	1
6	home	home	1
7	home	home	1
8	home	home	1
9	home	home	0.999
10	home	home	0.92
11	home	home	0.696
12	home	home	0.621
13	home	home	0.545
14	home	home	0.858
15	home	smb	0.64
16	smb	smb	0.6
17	smb	smb	0.971
18	smb	smb	0.984
19	smb	home	0.996
20	smb	smb	0.997
21	smb	smb	0.986

22	smb	smb	0.986
23	smb	smb	0.772
24	smb	smb	0.971
25	smb	smb	0.987
26	smb	smb	0.983
27	smb	smb	1
28	smb	smb	1
29	smb	smb	1
30	smb	smb	0.999
31	enterprise	enterprise	0.997
32	enterprise	enterprise	0.996
33	enterprise	enterprise	1
34	enterprise	enterprise	1
35	enterprise	enterprise	1
36	enterprise	enterprise	1
37	enterprise	enterprise	1
38	enterprise	enterprise	1
39	enterprise	enterprise	1
40	enterprise	enterprise	1
41	enterprise	enterprise	1
42	enterprise	enterprise	1
43	enterprise	enterprise	1
44	enterprise	enterprise	1
45	enterprise	enterprise	1

Perlu diketahui bahwa hasil perhitungan yang dilakukan bukanlah hasil akurasi dari keseluruhan klasifikasi yang dilakukan pada penelitian. Melainkan nilai hasil prediksi dari keseluruhan data yang digunakan. Sebagai contoh, data nomor 15 diklasifikasikan sebagai *label smb* karena memiliki nilai prediksi yang lebih besar ke arah *label smb* dengan nilai prediksi 0.640 atau 0.64, dan begitupun dengan data lain dengan nilai prediksinya masing-masing.

Untuk menghitung nilai akurasi dari keseluruhan data yang digunakan pada penelitian yang dilakukan ini, maka harus dilakukan berdasarkan nilai dari hasil *confusion matrix*.

Tabel 3. Tabel Hasil Confusion Matrix Pada Aplikasi Weka

Confusion Matrix	Pred Home	Pred SMB	Pred Enterprise
Act Home	14	1	0
Act SMB	1	14	0
Act Enterprise	0	0	15

Untuk menghitung nilai akurasi dari hasil *confusion matrix* dapat menggunakan persamaan rumus berikut:

$$A = \frac{TP}{N} \times 100\%$$

Keterangan:

- *A* : Nilai akurasi.
- *TP* : Nilai prediksi yang benar.
- *N* : Jumlah dokumen atau data yang digunakan.

Sehingga diperoleh hasil akurasi sebagai berikut:

$$A = \left(\frac{14 + 14 + 15}{45} \right) \times 100\% = 95,5556\%$$

Setelah dilakukan klasifikasi dari data yang sudah diproses pada tahap *pre-processing* data menggunakan aplikasi *Weka*, maka didapati bahwa data yang digunakan memiliki tingkat akurasi sebesar 95,5556% berdasarkan hasil perhitungan *confusion matrix* pada aplikasi *Weka*.

III. PEMBAHASAN

3.1 Pre-Processing Data

Data yang digunakan pada yang penelitian yang dilakukan ini merupakan data *pricelist* QNAP NAS yang didapat langsung dari PT. Distributor Trimitra Indonesia selaku pemilik nama dagang Toko *Storage*. Data awal pada *pricelist* QNAP NAS ini berjumlah 105 data produk yang berisikan 30 atribut data yang mencakup *segment*, *parts number*, *SRP*, hingga *max. number of concurrent connections (CIFS) - with max. memory*.

Sebelum data yang ada dapat digunakan untuk proses *mining* menggunakan algoritma *MNB*. Terlebih dahulu data diproses melalui beberapa tahapan pada tahap *pre-processing* data yang meliputi:

1. Seleksi Data

Pada tahap seleksi data ini dipilihlah 10 atribut dari total 30 atribut yang ada pada data *pricelist* produk QNAP. 10 atribut yang dipilih mencakup:

Tabel 4. Tabel Atribut Yang Dipilih Dari Data *Pricelist* QNAP

No	Nama Field
1	<i>Segment</i>
2	<i>SRP</i>
3	<i>Main Focus</i>
4	<i>Chassis</i>

5	<i>Drive Bay</i>
6	<i>M.2 Slot</i>
7	<i>CPU</i>
8	<i>Graphic Processors</i>
9	<i>System Memory</i>
10	<i>Max. Number of Concurrent Connections (CIFS) - with Max. Memory</i>

2. Transformasi Data

Tahap berikutnya adalah mentransformasi atau mengubah data yang sudah melewati tahap sebelumnya. Tahap transformasi data yang dilakukan mencakup:

- a) Perubahan nama atribut lama menjadi nama atribut baru.

Tabel 5. Tabel Contoh Perubahan Nama Atribut

No	Nama Atribut Lama	Nama Atribut Baru
1	<i>Segment</i>	<i>label</i>
2	<i>SRP</i>	<i>range_price</i>
3	<i>Main focus</i>	<i>main_focus</i>

- b) Perubahan nilai atribut lama menjadi nilai atribut baru.

Tabel 6. Tabel Contoh Perubahan Nilai Atribut

No	Nama Atribut	Nilai Atribut Baru
1	<i>label</i>	<i>home</i>
		<i>smb</i>
2	<i>range_price</i>	<i>enterprise</i>
		1jt sampai 20jt
		20jt sampai 40jt
		40jt sampai 60jt
		60jt sampai 80jt
3	<i>main_focus</i>	80jt sampai 100jt
		lebih dari 100jt
		<i>storage dan backup surveillance</i> atau <i>multimedia</i>

Contoh data yang nilai atribut lamanya diubah menjadi nilai atribut yang baru:

Tabel 7. Tabel Contoh Perubahan Nilai Atribut Lama

No	Nama Atribut	Nilai Atribut Lama	Nilai Atribut Baru
1	<i>label</i>	<i>Home Tower</i>	<i>home</i>
2	<i>range_price</i>	2.794.000	1jt sampai 20jt

3	<i>main_focus</i>	<i>Data Backup and Storage</i>	<i>storage dan backup</i>
---	-------------------	--------------------------------	---------------------------

c) Penggabungan nilai atribut menjadi satu atribut tunggal yang baru.

Tabel 8. Tabel Contoh Hasil Penggabungan Nilai Atribut

No	Kriteria	Label
1	1jt sampai 20jt <i>storage dan backup tower 2 bay</i> tidak <i>arm</i> tidak 2 <i>gb</i> 1 sampai 1000	<i>home</i>
2	20jt sampai 40jt <i>storage dan backup tower 9 bay</i> tidak <i>amd</i> tidak 32 <i>gb</i> 1001 sampai 2000	<i>smb</i>
3	20jt sampai 40jt <i>storage dan backup tower 8 bay</i> ya <i>intel</i> tidak 16 <i>gb</i> lebih dari 8000	<i>enterprise</i>

3. Pembersihan Data

Tahap berikutnya adalah melakukan pembersihan data yang mencakup:

- Menghapus data *label* selain *label* atau kelas *home*, *smb*, dan *enterprise*.
- Menghapus data yang bersifat duplikasi.
- Mengurangi jumlah data pada setiap *label* atau kelas agar data menjadi lebih seimbang dan tidak berpihak pada salah satu *label* atau kelas.

Hasil akhir dari proses ini didapatkan total data berjumlah 45 data dari 105 data awal yang ada, dengan masing-masing 15 data pada setiap *label*.

4. Tokenizing

Proses *tokenizing* dilakukan untuk memecah kalimat pada data atau dokumen menjadi *token* atau kata. Tujuan proses *tokenizing* ini adalah untuk memudahkan algoritma *MNB* dalam menghitung frekuensi kemunculan kata yang sama pada data yang digunakan. Berikut ini adalah contoh hasil *tokenizing* dari data yang digunakan:

Tabel 9. Tabel Contoh Hasil Proses Tokenizing

No	Spesifikasi	Label
1	['1jt', 'sampai', '20jt', 'storage', 'dan', 'backup', 'tower', '1', 'bay', 'tidak', 'arm', 'tidak', '2', 'gb', '1', 'sampai', '1000']	<i>home</i>

2	['1jt', 'sampai', '20jt', 'surveillance', 'atau', 'multimedia', 'tower', '2', 'bay', 'opsional', 'intel', 'ya', '4', 'gb', '1001', 'sampai', '2000']	<i>smb</i>
3	['20jt', 'sampai', '40jt', 'storage', 'dan', 'backup', 'tower', '8', 'bay', 'ya', 'intel', 'tidak', '16', 'gb', 'lebih', 'dari', '8000']	<i>enterprise</i>

5. Integrasi Data

Tahap terakhir pada pemrosesan data yang dilakukan adalah mengintegrasikan data dengan cara menggabungkan data yang sudah diproses pada tahap sebelumnya ke dalam file ARFF untuk kebutuhan pengujian dan validasi menggunakan aplikasi *Weka*, serta ke dalam file *Excel* untuk perhitungan manual.

3.2 Mining

Setelah tahap *pre-processing* data selesai dilakukan, berikutnya dilakukan proses *mining* dengan menggunakan algoritma *MNB*. Pada penelitian ini juga dilakukan perhitungan secara manual dengan menggunakan aplikasi *Microsoft Excel* untuk mencari tahu *label* dan juga nilai prediksi dari data sampel berikut ini:

Tabel 10. Tabel Data Sampel Hitungan Manual

No	Kriteria	Label
4	['1jt', 'sampai', '20jt', 'storage', 'dan', 'backup', 'tower', '2', 'bay', 'tidak', 'annapurnalabs', 'tidak', '1', 'gb', '1', 'sampai', '1000']	?

Tahap *mining* yang dilakukan meliputi:

1. Menghitung Probabilitas Label Terhadap Dokumen

Pada penelitian ini, total jumlah dokumen/data yang digunakan ada 45 data kriteria produk QNAP NAS yang terbagi menjadi 3 *label*, antara lain *label home* 15 data, *label smb* 15 data, dan *label enterprise* 15 data. Sehingga didapatkan hasil probabilitas dari setiap *label* adalah sebagai berikut:

Tabel 11. Tabel Probabilitas Setiap Label

No	Probabilitas Label
----	--------------------

1	$home = \frac{15}{45} = 0,33333$
2	$smb = \frac{15}{45} = 0,33333$
3	$enterprise = \frac{15}{45} = 0,33333$

2. Membuat Term-document Matrix

Langkah berikutnya adalah membuat *term-document matrix* yang berisikan jumlah kata pada seluruh dokumen, jumlah kata unik pada seluruh dokumen, serta jumlah kata pada setiap *label* dalam dokumen.

Tabel 12. Tabel Term-document Matrix

No	Keterangan	Jumlah Kata
1	Jumlah kata pada semua dokumen	777
2	Jumlah kata unik pada semua dokumen	52
3	Jumlah kata pada <i>label home</i>	255
4	Jumlah kata pada <i>label smb</i>	262
5	Jumlah kata pada <i>label enterprise</i>	260

3. Menghitung Probabilitas Kata Unik Pada Setiap Label

Langkah berikutnya adalah melakukan perhitungan probabilitas kata unik pada setiap *label* dengan menggunakan rumus yang sudah diterapkan teknik *laplace smoothing* untuk mencegah munculnya *zero probability*.

Tabel 13. Tabel Probabilitas Kata Unik

Kata	$\frac{Home (Kata) + 1}{255 + 52 }$	$\frac{SMB (Kata) + 1}{262 + 52 }$	$\frac{Enterprise (Kata) + 1}{260 + 52 }$
1jt	0,0521	0,0223	0,0032
sampai	0,1010	0,0987	0,0321
20jt	0,0521	0,0414	0,0064
...
1000	0,0326	0,0127	0,0032

4. Menghitung Probabilitas Dokumen Terhadap Label

Langkah terakhir yang harus dilakukan adalah menghitung nilai probabilitas dokumen atau data terhadap *label*. Berikut adalah perhitungan probabilitas dari data sampel yang digunakan:

$$- P(home|d(4)) = 0,3333 \times 0,0521 \times 0,1010 \times 0,0521 \times 0,0326 \times 0,0326 \times 0,0326 \times 0,0489 \times 0,0358 \times 0,0521 \times 0,0749 \times 0,0261 \times 0,0749 \times 0,0489 \times 0,0521 \times 0,0489 \times 0,1010 \times 0,0326 = 1,72561e^{-23}$$

$$- P(smb|d(4)) = 0,3333 \times 0,0223 \times 0,0987 \times 0,0414 \times 0,0287 \times 0,0287 \times 0,0287 \times 0,0287 \times 0,0064 \times 0,0510 \times 0,0382 \times 0,0096 \times 0,0382 \times 0,0127 \times 0,0510 \times 0,0127 \times 0,0987 \times 0,0127 = 9,6548e^{-28}$$

$$- P(enterprise|d(4)) = 0,3333 \times 0,0032 \times 0,0321 \times 0,0064 \times 0,0481 \times 0,0481 \times 0,0481 \times 0,0481 \times 0,0032 \times 0,0513 \times 0,0865 \times 0,0032 \times 0,0865 \times 0,0032 \times 0,0513 \times 0,0032 \times 0,0321 \times 0,0032 = 2,50388e^{-31}$$

Dari hasil perhitungan manual yang dilakukan pada 3 data sampel yang digunakan, didapati bahwa, data sampel nomor 4 masuk ke dalam *label home* karena memiliki nilai prediksi yang lebih tinggi dari *label* lainnya dengan nilai prediksi sebesar $1,72561e^{-23}$.

3.3 Pengujian Dan Validasi

Untuk membuktikan dan memvalidasi hasil perhitungan manual yang dilakukan, serta melakukan klasifikasi pada data kriteria yang lainnya. Maka, akan dilakukan juga tahap pengujian dan validasi menggunakan aplikasi *Weka*.

Berikut ini adalah hasil evaluasi dari perhitungan manual pada aplikasi *Weka* untuk membuktikan apakah hasil klasifikasi dari data sampel yang digunakan terbukti valid atau tidaknya. Perhitungan ini dilakukan dengan menggunakan persamaan rumus sebagai berikut:

$$np(c) = \frac{P(c|d(n))}{\sum P(c|d(n))}$$

Keterangan:

- $np(c)$: Nilai prediksi dari *label*.
- $P(c|d(n))$: Probabilitas kalimat terhadap *label*.

- $\sum P(c|d(n))$: Jumlah probabilitas dari semua *label*.

Sehingga didapat hasil akhir sebagai berikut untuk data sampel nomor 4:

- $np(home) = 1,72561e^{-23} / (1,72561e^{-23} + 9,6548e^{-28} + 2,50388e^{-31}) = 1,72561e^{-23} / 1,7257e^{-23} = 1,000$
- $np(smb) = 9,6548e^{-28} / (1,72561e^{-23} + 9,6548e^{-28} + 2,50388e^{-31}) = 9,6548e^{-28} / 1,7257e^{-23} = 0,000$
- $np(enterprise) = 2,50388e^{-31} / (1,72561e^{-23} + 9,6548e^{-28} + 2,50388e^{-31}) = 2,50388e^{-31} / 1,7257e^{-23} = 0,000$

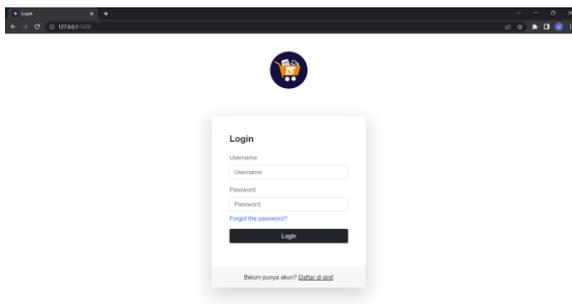
Setelah mendapatkan hasil nilai prediksi dari semua *label* yang ada, langkah berikutnya adalah membandingkan hasil nilai prediksi dari setiap *label* untuk menemukan nilai prediksi tertinggi. Dari hasil perhitungan yang sudah dilakukan didapatkan bahwa, Data sampel nomor 4 diklasifikasikan sebagai *label home* karena memiliki nilai prediksi tertinggi yaitu sebesar **1,000** dibandingkan nilai prediksi dari *label* lain.

Dari hasil tersebut dapat disimpulkan bahwa perhitungan manual yang dilakukan terbukti valid karena menghasilkan klasifikasi yang tepat dan sesuai dengan hasil klasifikasi pada aplikasi *Weka* yang dapat dilihat pada **Tabel 2**.

3.4 Tampilan Aplikasi

1. Halaman *Login*

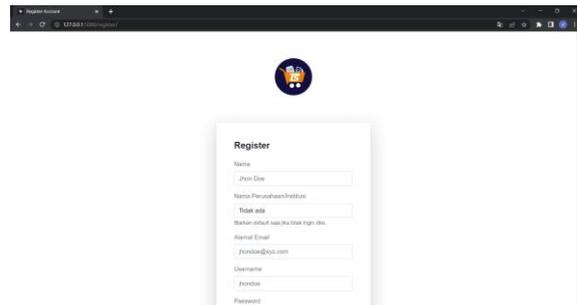
Halaman *login* dapat digunakan oleh pengguna untuk masuk dan menggunakan sistem rekomendasi yang telah dibuat.



Gambar 1. Gambar Halaman *Login*

2. Halaman *Register*

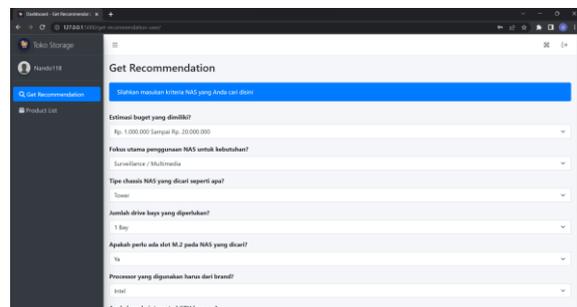
Halaman *register* dapat digunakan oleh pengguna untuk membuat akun agar dapat menggunakan sistem yang telah dibuat.



Gambar 2. Gambar Halaman *Register*

3. Halaman *Get Recommendation*

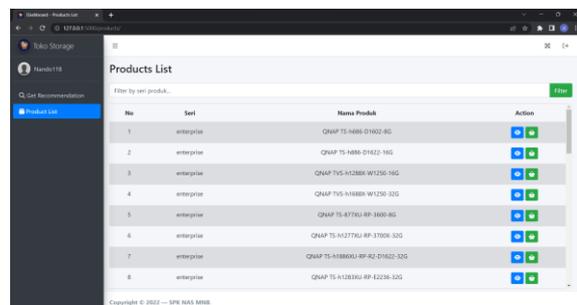
Halaman *get recommendation* dapat digunakan oleh pengguna untuk mendapatkan rekomendasi seri produk *NAS*.



Gambar 3. Gambar Halaman *Get Recommendation*

4. Halaman *Products List*

Halaman *products list* berisikan daftar produk *QNAP NAS* yang sudah dimasukkan oleh *admin* ke dalam sistem rekomendasi yang telah dibuat.



Gambar 4. Gambar Halaman *Products List*

3.5 Pengujian Aplikasi

Untuk mengetahui apakah fungsionalitas yang sudah dibuat dapat berjalan dengan semestinya sewaktu diterapkan data asli, maka dilakukan pengujian aplikasi

menggunakan metode *black box testing*. Adapun hasil pengujiannya adalah sebagai berikut.

1. Pengujian Halaman Login

Tabel 14. Tabel Pengujian Black Box Modul Login

No	Skenario Pengujian	Hasil Yang Diharapkan	Hasil Pengujian	Keterangan
1	Login sebagai admin	Berhasil login dan menampilkan halaman admin.	Sesuai	Valid
2	Login sebagai konsumen	Berhasil login dan menampilkan halaman konsumen.	Sesuai	Valid
3	Login dengan akun yang belum terdaftar	Tidak berhasil login dan menampilkan pesan akun tidak terdaftar.	Sesuai	Valid
4	Login dengan password yang salah	Tidak berhasil login dan menampilkan pesan password salah.	Sesuai	Valid
5	Logout dari halaman admin	Berhasil logout dan menampilkan halaman login.	Sesuai	Valid
6	Logout dari halaman konsumen	Berhasil logout dan menampilkan halaman login.	Sesuai	Valid

2. Pengujian Halaman Register

Tabel 15. Tabel Pengujian Black Box Modul Register

No	Skenario Pengujian	Hasil Yang Diharapkan	Hasil Pengujian	Keterangan
1	Membuat akun baru	Berhasil dan menampilkan halaman login serta pesan	Sesuai	Valid

2	Membuat akun dengan username yang sudah terdaftar Login dengan akun yang baru dibuat	Tidak berhasil dan menampilkan pesan <i>username</i> sudah terdaftar. Berhasil login dan menampilkan halaman konsumen.	Sesuai	Valid
3			Sesuai	Valid

3. Pengujian Halaman Get Recommendation

Tabel 16. Tabel Pengujian Black Box Modul Get Recommendation

No	Skenario Pengujian	Hasil Yang Diharapkan	Hasil Pengujian	Keterangan
1	Input kriteria NAS yang dicari	Berhasil dan menampilkan pesan rekomendasi seri produk NAS serta menampilkan tabel produk NAS dengan seri yang direkomendasikan.	Sesuai	Valid
2	Input kriteria NAS dari data sampel nomor 4	Berhasil dan menampilkan pesan rekomendasi seri home serta menampilkan tabel produk NAS dengan seri home.	Sesuai	Valid
3	Input kriteria NAS dari data sampel nomor 17	Berhasil dan menampilkan pesan rekomendasi seri smb serta menampilkan tabel produk NAS dengan seri smb.	Sesuai	Valid
4	Input kriteria NAS dari data	Berhasil dan menampilkan pesan rekomendasi	Sesuai	Valid

sampel nomor 33	seri <i>enterprise</i> serta menampilkan tabel produk <i>NAS</i> dengan seri <i>enterprise</i> .
-----------------	--

4. Pengujian Halaman *Products List*

Tabel 17. Tabel Pengujian *Black Box* Halaman *Products List*

No	Skenario Pengujian	Hasil Yang Diharapkan	Hasil Pengujian	Keterangan
1	Menekan tombol <i>action check product</i>	Berhasil dan menampilkan halaman produk <i>NAS</i> pada <i>website</i> resmi QNAP.	Sesuai	Valid
2	Menekan tombol <i>buy product</i>	Berhasil dan menampilkan halaman pembelian produk.	Sesuai	Valid
3	<i>Filter</i> produk <i>NAS</i> berdasarkan seri <i>smb</i>	Berhasil dan menampilkan data produk <i>NAS</i> dengan seri <i>smb</i> .	Sesuai	Valid

dapat diakses dengan mudah oleh pengguna untuk mendapatkan rekomendasi seri produk *NAS* berdasarkan hasil kuesioner pengujian sistem.

4.2 Saran

Dari hasil penelitian dan analisa yang telah dilakukan, penulis menyadari bahwa penelitian yang dilakukan masih jauh dari kata sempurna. Oleh sebab itu diberikan beberapa saran agar penelitian ini dapat lebih baik lagi. Adapaun saran yang diberikan adalah sebagai berikut:

Diharapkan agar penelitian yang selanjutnya dapat menggunakan data yang lebih baik lagi agar proses pengolahan data tidak terlalu Panjang; Diharapkan agar penelitian yang selanjutnya dapat menggunakan algoritma atau metode lain untuk membandingkan hasilnya; Diharapkan agar penelitian yang selanjutnya dapat mengembangkan sistem menjadi *mobile app* dan memperbaiki kekurangan pada *UI/UX* dari sistem yang dibuat.

IV. KESIMPULAN DAN SARAN

4.1 Kesimpulan

Dari hasil pengujian yang telah dilakukan pada penelitian ini, maka dapat diambil kesimpulan sebagai berikut:

Penerapan metode *data mining* dan algoritma *multinomial naïve bayes* dapat diimplementasikan untuk merekomendasikan seri produk *NAS* berdasarkan kriteria yang dicari oleh calon konsumen. Hal ini dibuktikan dari hasil pengujian dan evaluasi yang menghasilkan nilai akurasi sebesar 95,5556% dari data yang digunakan pada penelitian ini; Sistem yang dirancang terbukti

DAFTAR PUSTAKA

- [1] R. Rinaldi, R. Goejantoro, dan Syaripuddin, “Penerapan Metode Klasifikasi Multinomial Naive Bayes (Studi Kasus: PT Prudential Life Samarinda Tahun 2019),” *EKSPONENSIAL*, vol. 12, no. 2, hal. 111–118, 2021, [Daring]. Tersedia pada: <https://repository.unmul.ac.id/handle/123456789/13229>.
- [2] S. Wahyuni dan Adinda, “Penerapan Algoritma Naïve Bayes Pada Prediksi Kelulusan Mahasiswa (Studi Kasus: Institut Medika Drg. Suherman),” *SIMANTIK*, vol. 6, no. 2, hal. 29–34, 2021, [Daring]. Tersedia pada: <https://simantik.panca-sakti.ac.id/index.php/simantik/article/view/120/116>.
- [3] W. K. Cikutra, S. Hariyanto, E. D. Kusuma, dan R. Rimbawan, “Identifikasi Penyalahgunaan Jalur SMS Marketing Dengan Konten SMS Premium Pada PT Dian Prima Jayaraya Dengan Algoritma Naive Bayes,” *ALGOR*, vol. 3, no. 1, hal. 14–26, 2021, doi: <https://doi.org/10.31253/algor.v3i1.658>.
- [4] E. A. Firdaus, S. Maulani, dan A. B. Dharmawan, “Pengukuran Minat Baca Mahasiswa Dengan Metode Clustering Di Perpustakaan Akademi Keperawatan RS.Dustira Cimahi Menggunakan Data Mining,” *NUANSA Inform.*, vol. 15, no. 1, hal. 32–40, 2021, doi: <https://doi.org/10.25134/nuansa>.
- [5] A. Jananto, S. Sulastri, E. Nur Wahyudi, dan S. Sunardi, “Data Induk Mahasiswa sebagai Prediktor Ketepatan Waktu Lulus Menggunakan Algoritma CART Klasifikasi Data Mining,” *SISFOKOM*, vol. 10, no. 1, hal. 71–78, Feb 2021, doi: [10.32736/sisfokom.v10i1.991](https://doi.org/10.32736/sisfokom.v10i1.991).
- [6] F. E. Purwiantono dan A. Aditya, “Klasifikasi Sentimen Sara, Hoaks Dan Radikal Pada Postingan Media Sosial Menggunakan Algoritma Naive Bayes Multinomial Text,” *TEKNOKOMPAK*, vol. 14, no. 2, hal. 68–73, Agu 2020, doi: [10.33365/jtk.v14i2.709](https://doi.org/10.33365/jtk.v14i2.709).
- [7] C. N. Harahap, G. I. Marthasari, dan N. Hayatin, “Perbandingan Klasifikasi Berita Hoax Kategori Kesehatan Menggunakan Naïve Bayes dan Multinomial Naïve Bayes,” *REPOSITOR*, vol. 3, no. 4, hal. 419–424, 2021, doi: <https://doi.org/10.22219/repositor.v3i4.1380>.
- [8] Ainurrohmah, “Akurasi Algoritma Klasifikasi pada Software Rapidminer dan Weka,” *PRISMA*, vol. 4, no. 1, hal. 493–499, 2021, [Daring]. Tersedia pada: <https://journal.unnes.ac.id/sju/index.php/prisma/article/view/45090/18550>.
- [9] M. Rahmadi, F. Kaurie, dan T. Susanti, “Uji Akurasi Dataset Pasien Pasca Operasi Menggunakan Algoritma Naïve Bayes Menggunakan Weka Tools,” *JURIKOM (Jurnal Ris. Komputer)*, vol. 7, no. 1, hal. 134–139, Feb 2020, doi: [10.30865/jurikom.v7i1.1761](https://doi.org/10.30865/jurikom.v7i1.1761).
- [10] R. K. Ngantung dan M. A. I. Pakereng, “Model Pengembangan Sistem Informasi Akademik Berbasis User Centered Design Menerapkan Framework Flask Python,” *MEDIA Inform. BUDIDARMA*, vol. 5, no. 3, hal. 1051–1062, Jul 2021, doi: [10.30865/mib.v5i3.3054](https://doi.org/10.30865/mib.v5i3.3054).

BIOGRAFI

Fernando Verdy Sunata, lahir di Tangerang pada tanggal 18 Januari 2000. Menyelesaikan pendidikan Strata I (S1) pada tahun 2022 pada Program Studi Teknik Informatika di Universitas Buddhi Dharma. Saat ini bekerja sebagai *staff Admin* di PT. Distributor Trimitra Indonesia.

Susanto Hariyanto, lahir di Pontianak pada tahun 1986. Menyelesaikan Magister Komputer di STMIK Eresha tahun 2012. Saat ini mengajar sebagai dosen tetap pada Program Studi Teknik Informatika di Universitas Buddhi Dharma sejak tahun 2019.

Hartana Wijaya, saat ini bekerja sebagai dosen Tetap pada Program Studi Teknik Informatika di Universitas Buddhi Dharma.