

Analisis Metode Klasifikasi Data Naïve Bayes dan SVM Dalam Menentukan Keunikan Hotel

Ahmad Afif¹

¹Universitas Adzkie Sumatera Barat

¹Program Studi Informatika, Universitas Adzkie Sumatera Barat

e-mail: ahmadafif@adzkie.ac.id

Abstract

At present, the national tourism sector became a new prima donna for national development. Foreign exchange contributions and employment in this sector are very significant for the country's foreign exchange. In fact, it is estimated that by 2019 it has defeated foreign exchange earnings from the palm oil industry (CPO). In this case, the government must increase the growth of tourist visits coming to Indonesia. One part that needs to be considered by the government in improving the tourism sector is hotel accommodation. In improving hotel accommodation services, there needs to be a service that contains information about the uniqueness of the hotel. The service to be developed has a two-way relationship technique between customers and service providers. This two-way relationship occurs by grouping hotel types based on comment data google. The main purpose of this paper is to analyze several methods that are in accordance with the classification of hotel uniqueness. The uniqueness of the hotel to be classified is a hotel that has natural, European, classic, photo and home themes. The method to be compared is the Support Vector Machine (SVM) and Naïve Bayes methods. The selection of the Naïve Bayes and SVM methods was carried out because these two methods have fairly good accuracy. In this research, it can be produced that the accuracy of Naïve Bayes is higher than the accuracy of SVM, the comparison is 75% and 62.5%.

Keywords: *Tourism, Naïve Bayes, SVM, Text Mining, Recommendation.*

Abstrak

Pada saat ini, sektor pariwisata nasional menjadi primadona baru bagi pembangunan nasional. Kontribusi devisa dan penyerapan tenaga kerja di sektor ini sangat signifikan bagi devisa negara. Bahkan, diperkirakan pada 2019 sudah mengalahkan perolehan devisa dari industri sawit (CPHAI). Pada kasus ini, pemerintah harus meningkatkan pertumbuhan kunjungan wisatawan yang datang ke Indonesia. Salah satu bagian yang perlu diperhatikan oleh pemerintah dalam meningkatkan sektor pariwisata adalah akomodasi hotel. Dalam meningkatkan pelayanan akomodasi hotel, perlu adanya suatu pelayanan yang memuat informasi tentang keunikan hotel tersebut. Layanan yang akan dikembangkan memiliki teknik hubungan dua arah antara pelanggan dan penyedia layanan. Hubungan dua arah ini terjadi dengan mengelompokkan jenis hotel berdasarkan data komentar di google. Tujuan utama dari penelitian ini adalah menganalisis beberapa metode yang sesuai dengan klasifikasi keunikan hotel. Keunikan hotel yang akan diklasifikasikan adalah hotel yang bertema alam, eropa, klasik, foto dan nuansa rumah. Metode yang akan dibandingkan adalah metode Support Vector Machine (SVM) dan metode Naïve Bayes. Pemilihan metode Naïve Bayes dan SVM dilakukan karena kedua metode ini memiliki akurasi yang cukup baik. Pada penelitian ini dapat dihasilkan bahwa akurasi Naïve Bayes lebih tinggi dibandingkan dengan akurasi SVM, perbandingannya adalah 75% dan 62,5%.

Kata Kunci: Pariwisata, Naïve Bayes, SVM, Text Mining, Rekomendasi.

1. PENDAHULUAN

Pada saat ini, sektor pariwisata Indonesia memiliki peran penting dalam pendapatan keuangan negara. Hal ini disebabkan turunnya nilai ekspor Indonesia, baik dari sisi migas maupun nonmigas. Pada tahun 2018, Pemerintah menargetkan kunjungan wisatawan ke Indonesia sebanyak 10 juta dengan tolok ukur perolehan devisa sebesar Rp 142 triliun. Selain itu, dalam lima tahun ke depan, pemerintah menetapkan target sebanyak 20 juta wisman,

dengan target sebesar Rp 260 triliun. Angka tersebut sangat wajar, mengingat selama tahun 2014 sektor pariwisata Indonesia menyumbang devisa sebesar Rp136 triliun. Industri pariwisata Indonesia sudah memiliki pertumbuhan yang baik, yaitu sebesar 7,2 persen per tahun. Angka ini bahkan lebih tinggi dari rata-rata pertumbuhan pariwisata dunia yang hanya tercatat 4,7 persen. Dengan jumlah wisatawan dunia yang mencapai 1,3 miliar orang, masih terdapat potensi untuk meningkatkan pertumbuhan kunjungan wisatawan.

Untuk meningkatkan pertumbuhan kunjungan wisatawan, pemerintah harus mengambil langkah konkrit dan lebih serius di bidang pariwisata. Langkah-langkah tersebut antara lain meningkatkan pengembangan industri pariwisata Indonesia. Peningkatan pengembangan industri pariwisata dapat dilakukan dengan menggunakan media teknologi dalam pelaksanaannya. Sektor pariwisata secara substansial diuntungkan oleh teknologi, terutama dalam hal teknologi internet [2]. Indonesia tergolong negara kepulauan yang memiliki wilayah geografis luas dan kebudayaan yang sangat beragam. Keberagaman budaya ini bisa menjadi salah satu faktor penarik wisatawan untuk datang ke Indonesia.

Media internet telah menjadi sumber informasi yang sangat penting dan sangat dibutuhkan oleh wisatawan dalam berwisata, seperti kapan harus memilih destinasi wisata, media transportasi, destinasi kuliner, akomodasi hotel dan lain-lain. Pencarian informasi secara online akan mendapat respon yang baik dari wisatawan pada saat merencanakan perjalanan, reservasi hotel dan pembelian tiket [3]. Saat ini, wisatawan dan penyedia jasa pariwisata dapat mencari, membandingkan dan memilih destinasi wisata yang diinginkan. Pada tahun 2014, sebanyak 30% pesan dilakukan secara online, dan persentase ini diperkirakan akan terus meningkat berkali-kali lipat selama beberapa tahun ke depan [2]. Pemesanan yang dilakukan wisatawan berupa pemesanan tiket pesawat, wahana wisata, pemesanan hotel dan lain-lain.

Salah satu bagian yang perlu diperhatikan oleh pemerintah dalam meningkatkan sektor pariwisata adalah akomodasi hotel. Setiap wisatawan pasti membutuhkan tempat untuk beristirahat dan melepas rasa lelah saat berwisata. Dengan meningkatkan pelayanan akomodasi hotel maka wisatawan, maka wisatawan akan lebih nyaman dalam berwisata dan akan meningkatkan jumlah wisatawan yang datang ke Indonesia. Dalam meningkatkan pelayanan akomodasi hotel, perlu adanya suatu pelayanan yang memuat informasi tentang keunikan hotel tersebut. Keunikan hotel yang akan diklasifikasikan adalah hotel yang memiliki tema atau keunikan tertentu. Tema atau keunikan hotel tersebut dapat berupa tema alam, eropa, klasik, foto dan rumah. Layanan ini akan berbasis Android. Layanan Android ini dipilih karena saat ini layanan Android sudah banyak digunakan oleh masyarakat, termasuk wisatawan. Hal ini dilakukan agar wisatawan dapat mengakses layanan ini dimana saja dan kapan saja. Layanan ini akan mengklasifikasikan jenis hotel dengan tema tertentu. Tema hotel ini adalah ciri khas, suasana dan fasilitas hotel. Tema-tema ini termasuk tema alam, modern, klasik, rumah dan foto. Dengan layanan ini, wisatawan akan dapat memilih hotel berdasarkan minat, hobi, dan karakter wisatawan.

Layanan yang akan dikembangkan memiliki teknik hubungan dua arah antara pelanggan dan penyedia layanan. Hubungan dua arah ini terjadi dengan mengelompokkan jenis hotel berdasarkan data komentar google. Hasil dari layanan yang diproses akan memberikan info baru bagi wisatawan tentang klasifikasi keunikan hotel. Data yang diolah adalah data komentar wisatawan tentang keunikan hotel tersebut. Data komentar ini merupakan data berupa teks. Untuk mengambil pengetahuan dari data teks diperlukan cara yang baik dan efisien. Salah satu cara yang baik dan efisien untuk mengekstraksi pengetahuan dari data teks adalah dengan *text mining*.

Salah satu proses dalam *text mining* adalah proses klasifikasi. Klasifikasi adalah suatu fungsionalitas data mining yang menghasilkan model untuk memprediksi kelas atau kategori dari objek - objek didalam basis data. Klasifikasi merupakan proses yang terdiri dari dua tahap, yaitu tahap pembelajaran dan tahap pengklasifikasian. Proses klasifikasi *text mining* dapat

berupa klasifikasi sentimen (positif dan negatif) dan klasifikasi tipe data. Pada penelitian ini akan dilakukan proses *text mining* dengan mengklasifikasikan jenis data teks. Data teks akan diklasifikasikan menurut keunikan suatu hotel. *Text Mining* adalah penambangan data berupa teks, yang memiliki tahapan proses yang lebih banyak karena data teks. Teks merupakan data yang memiliki informasi yang tidak terstruktur sehingga tahapan prosesnya lebih banyak dengan tujuan agar data menjadi lebih terstruktur yang dapat diperoleh melalui pola-pola yang ada pada teks. Tujuan utama dari makalah ini adalah menganalisis beberapa metode yang sesuai dengan klasifikasi keunikan hotel. Metode yang akan dibandingkan adalah metode Support Vector Machine (SVM) dan metode Naïve Bayes.

2. METODE PENELITIAN

A. Kumpulan Data Penelitian

Dataset yang digunakan dalam penelitian ini menggunakan data komentar dari teks komentar di Google Map. Data diambil secara manual dari Google Maps. Data yang telah diambil akan dimasukkan ke dalam database dan nantinya menjadi dataset dalam pengolahan data. Data yang akan diolah berupa teks dan telah ditentukan kategorinya secara manual yaitu hotel dengan tema alam, eropa, klasik, foto dan rumah. Kumpulan data yang akan diproses berjumlah 120 komentar. Dari 120 data komentar, 100 dataset komentar akan digunakan sebagai data pelatihan dan 20 dataset komentar akan digunakan sebagai data pengujian.

Pra-pemrosesan Data

- Transformasi Kasus

Tahapan ini merupakan proses mengubah bentuk semua kata dalam dataset penelitian. Perubahan ini membuat semua huruf dalam dataset penelitian menjadi huruf kecil.

- Tokenisasi

Tahap kedua ini merupakan tahap pemisahan setiap kata dalam dataset penelitian. Hasil pemisahan data teks disebut token

- Penyaringan

Pada tahap ini, kata-kata yang sering muncul dan tidak berpengaruh pada klasifikasi teks akan dihilangkan. Kata yang termasuk adalah petunjuk waktu, kata tanya, kata sambung dan lain-lain.

B. Seleksi Fitur

Salah satu fitur dalam proses klasifikasi adalah fitur N-gram. Fitur N-gram digunakan dalam proses pemodelan dengan cara membagi sebuah kalimat menjadi beberapa bagian kata. Dalam N-gram, 'N' menunjukkan jumlah kata yang akan dikelompokkan menjadi satu bagian. Terdapat beberapa fitur N-gram yang telah dikembangkan yaitu fitur unigram yang membagi kalimat menjadi satu bagian, fitur bigram yang akan membagi kalimat menjadi dua bagian dan fitur trigram yang membagi kalimat menjadi tiga bagian. Dalam mengklasifikasikan data, penelitian ini menggunakan fitur unigram.

C. SVM

SVM menjadi salah satu metode yang baik dalam proses regresi dan klasifikasi data. SVM dapat bekerja dalam dataset dengan jumlah besar yang dapat diambil dari situs online dan menjadi populer karena penerapannya dalam proses klasifikasi teks, disambiguasi kata sensorik. SVM merupakan salah satu strategi pembelajaran yang dapat digunakan sebagai metode klasifikasi dengan mempertimbangkan kapasitas polinomial, kapasitas premis, dan sistem saraf. Untuk masalah yang tidak dapat langsung diisolasi dalam ruang informasi, SVM menawarkan kemungkinan untuk menemukan jawaban dengan membuat peningkatan ruang informasi utama menjadi ruang fitur berdimensi tinggi [9].

D. Naive Bayes

Naïve Bayes adalah proses klasifikasi berdasarkan Hipotesis Bayes. Dalam proses klasifikasi ini, Hipotesis Bayes akan menentukan hubungan antara kemungkinan dua peluang A dan B, P(A) dan P(B) dan kemungkinan terjadinya kontinjensi A yang dicetak oleh B dan peluang B diadaptasi oleh A. $P(A | B)$ dan $P(B | A)$. Persamaan Bayes dapat dilihat pada rumus di bawah ini:

$$P(A/B) = \frac{P(B/A)P(A)}{P(B)}$$

Naïve Bayes dapat mengukur kemungkinan suatu peristiwa dengan mempertimbangkan kasusnya. Sejalan dengan itu, evaluasi terhadap kemungkinan suatu peristiwa dapat dilakukan dengan mempertimbangkan kasus sebelumnya.

3. HASIL DAN PEMBAHASAN

A. Transformasi Kasus

Dengan menggunakan pustaka *case folding* di python, data komentar akan terbentuk dengan semua teks huruf kecil. Hasil dari kasus transformasi data dapat dilihat pada tabel di bawah ini.

Tabel 1. Tabel Data Komentar

Komentar
Hotel nyaman. Ada fasilitas kolam renang. Konsep ala2 Jawa jaman dulu. Sedang di bangun pendopo untuk menyambut tamunya. Harga terjangkau. Lokasi di tengah kota

Tabel 1 memperlihatkan salah satu data komentar yang telah diambil. Dari data tersebut dapat diambil kata kunci dalam proses klasifikasi. Kata-kata "jawa", "jaman", "dulu" dan "pendopo", bisa menjadi kata kunci untuk dikategorikan sebagai tema klasik nantinya.

B. Tokenisasi

Dengan menggunakan *library tokenized* di python, data komentar akan dibentuk secara terpisah token (token). Hasil tokenisasi dapat dilihat pada Gambar 1.

```
Out[38]: ['aahh',  
         'ac',  
         'acara',  
         'ada',  
         'adat',  
         'adem',  
         'agak',  
         'air',  
         'akan',  
         'akses',  
         'aku',  
         'ala',  
         'alam',  
         'alami',  
         'alkid',  
         'amat',  
         'anda',  
         'antara',  
         'antik',
```

Gambar 1. Hasil *Tokenisasi*

C. Penyaringan

Dengan menggunakan *library filtering* di python, data komentar akan terbentuk dengan data tanpa kata-kata umum yang tidak penting. Hasil penyaringan dapat dilihat pada Gambar 2.

0	suasana kental adat jawa namun untuk sarapan p...	klasik
1	hotel nyaman ada fasilitas kolam renang konsep...	klasik
2	hotel yang menyuguhkan suasana yang nyaman etn...	klasik
3	benar seperti di rumah sendiri kental akan sua...	rumah
4	nginep serasa di rumah orang tua sendiri pelay...	rumah
5	kolam renang sepi enak serasa rumah sendiri deh	rumah
6	penginapan bergaya trradisional yang menarik n...	klasik
7	pertama kali masuk ke guesthouse ini saya suda...	klasik
8	tradisional agak serem tapi sebenarnya nyaman sih	klasik
9	bangunan tampak klasik dengan model dinding ta...	klasik
10	penginapan dengan suasana jawa deket dengan pu...	klasik
11	menginap malam di hotel ini sangat nyaman deng...	klasik
12	ikut pelatihan hiperkes disini desain khas hot...	klasik
13	untuk yang budget tipis penginapan ini cocok b...	klasik

Gambar 2. Hasil penyaringan

D. Pelatihan Data

Dari 120 kata dalam dataset penelitian. Sebanyak 100 kata akan digunakan sebagai data pelatihan. Pada gambar akan ditampilkan data *training* yang akan digunakan.

```
In [16]: X_train
Out[16]: 17 tempatnya bagus banget cocok buat refreshing k...
          22 berada di parangtritis yang mempunyai pemandan...
          30 tempat yang sejuk sepi dari keramaian dan pema...
          63 bagus banget pelayanan ok kolam renang dan desa...
           4 nginep serasa di rumah orang tua sendiri pelay...
           3 benar seperti di rumah sendiri kental akan sua...
          21 tempat yang nyaman pelayanan ramah suasana asr...
          40 penginapan model rumah jawa ac dan bed ready a...
          16 hotel nya bersih pelayanan bagus pemandangan b...
          45 lokasinya tenang dan cukup nyaman jauh dari hi...
           9 bangunan tampak klasik dengan model dinding ta...
          18 terpatnya jawa banget sesuai dengan namanya jo...
          32 interior mulai dari lobby dan ruangan ruangan...
          46 tempatnya sunyi sejuk dan nyaman ada kolam ren...
```

Gambar 3. Data *Training Pelatihan Data*

E. Pengujian Data

Dari 120 kata dalam dataset penelitian, sebanyak 20 kata akan digunakan sebagai data pengujian. Pada gambar akan ditampilkan data training yang akan digunakan tersebut.

```
In [17]: X_test
Out[17]: 38 suasana nya seperti tinggal dirumah jaman dulu...
          11 menginap malam di hotel ini sangat nyaman deng...
          20 pertama kali kesini langsung suka tempatnya ad...
          26 rumah adat jawa dengan pelayanan yang baik
          31 penginapan di sleman jogja tepatnya di daerah ...
          23 tempat indah apalagi kalau dapat sunset jalan ...
          42 tempat yang cocok untuk mencari ketenangan ja...
          66 tempat yang cocok untuk menginap di jogja bers...
```

Gambar 4. Training pengujian data

F. Penomoran Klasifikasi

Pada proses ini, semua tema yang telah diklasifikasi akan diberi label berupa angka. Proses ini dilakukan untuk memudahkan pengklasifikasian. Dalam pemberian label, tema "Alam" diberi label "0", tema "Klasik" diberi label "1", tema "Rumah" diberi label "2", tema "Foto" diberi label "3" dan tema "Eropa" diberi label "4".

```
In [6]: komen.klasifikasi.value_counts()
Out[6]: Rumah      24
        Alam       24
        Eropa      24
        Klasik     24
        Foto       24
        Name: klasifikasi, dtype: int64
```

Gambar 5. Pelabelan data

G. Hasil Akurasi

Dengan menggunakan library sklearn yang telah terdapat pada pemrograman python, dapat dihasilkan klasifikasi Naive bayes dan SVM seperti pada Gambar 6 dan Gambar 7. Pada setiap citra terdapat data komentar yang diberi nomor dan juga klasifikasi data yang dikomentari dengan tipe tertentu. Tema-tema ini telah diberi label dengan angka 0 sampai 4.

```

In [23]: X_test
Out[23]: 13 Eco friendly hotel in yogyakarta. Konsepnya ba...
          8 Tempat yang pas untuk kegiatan outbond acara k...
          15 Resort terbaik yg pernah sy kunjungi. Byk spo...
          94 tempat mantap dengan desain ala Eropa khususny...
          113 Pokoknya yang bikin berkesan tuh straf nya yan...
          73 Hotelnya sangat nyaman, lokasi juga strategis....
          107 Tempat yang nyaman dan kental traditional, rek...
          110 Suasana nyaman, seperti diRumah nenek
          118 hotelnya bagus, kayak di paris
          83 Keren sih. Paris banget. Cuma ya mahal kalo ma...
          70 Instagrammable hotel, buat yang senang Foto2. ...
          12 Resort ini menawarkan konsep danau dan gezebo ...
          26 Nuansa kampungnya dapet

In [25]: y_test
Out[25]: 13 0
          8 0
          15 0
          94 4
          113 3
          73 3
          107 1
          110 2
          118 4
          83 4
          70 3
          12 0
          26 1
    
```

Gambar 6. Klasifikasi Naïve Bayes

```

In [36]: X_test
Out[36]: 13 Eco friendly hotel in yogyakarta. Konsepnya ba...
          8 Tempat yang pas untuk kegiatan outbond acara k...
          15 Resort terbaik yg pernah sy kunjungi. Byk spo...
          94 tempat mantap dengan desain ala Eropa khususny...
          113 Pokoknya yang bikin berkesan tuh straf nya yan...
          73 Hotelnya sangat nyaman, lokasi juga strategis....
          107 Tempat yang nyaman dan kental traditional, rek...
          110 Suasana nyaman, seperti diRumah nenek
          118 hotelnya bagus, kayak di paris
          83 Keren sih. Paris banget. Cuma ya mahal kalo ma...
          70 Instagrammable hotel, buat yang senang Foto2. ...
          12 Resort ini menawarkan konsep danau dan gezebo ...
          26 Nuansa kampungnya dapet

In [34]: y_test
Out[34]: 13 0
          8 0
          15 0
          94 4
          113 3
          73 3
          107 1
          110 2
          118 4
          83 4
          70 3
          12 0
          26 1
    
```

Gambar 7. Klasifikasi SVM

Proses klasifikasi data komentar hotel dilakukan dengan menggunakan python sebagai alat klasifikasi. Dalam proses klasifikasi dan pengukuran akurasi, penelitian ini menggunakan sklearn library. Setelah dilakukan pengujian, hasilnya dapat dilihat pada Tabel 2.

Tabel 2. Hasil klasifikasi data

Metode	Akurasi
Naïve Bayes	0.75
SVM	0.625

4. PENUTUP

Penelitian ini bertujuan untuk menganalisis perbandingan akurasi antara Naïve Bayes dan SVM dalam proses klasifikasi data komentar hotel. Setelah dilakukan preprocessing data dan pemilihan fitur, dapat disimpulkan bahwa dalam pengklasifikasian hotel berdasarkan keunikan hotel, metode Naïve Bayes memiliki akurasi yang lebih tinggi dibandingkan dengan metode SVM. Metode Naïve Bayes memiliki akurasi sebesar 75% sedangkan SVM memiliki akurasi sebesar 62,5%. Berdasarkan hasil klasifikasi tersebut, peneliti akan menggunakan metode Naive Bayes dalam mengklasifikasikan data jasa akomodasi hotel. Layanan ini akan mendukung wisatawan dalam memilih hotel sesuai dengan tema yang mereka sukai dan layanan ini akan berjalan pada platform android.

Penelitian ini masih memiliki beberapa kelemahan dan kendala, antara lain sedikitnya jumlah data set yang digunakan dan sedikitnya jenis hotel yang diklasifikasikan. Masalah ini muncul karena ada beberapa rintangan dalam proses pengambilan data, di antara beberapa kendala tersebut adalah sulitnya mengambil data secara otomatis dari komentar di google maps. Dalam pengambilan datanya, Google Maps mengharuskan peneliti untuk melakukan pembayaran yang cukup mahal. Jadi, masalah ekonomi peneliti juga menjadi kendala dalam pengambilan data komentar. Untuk penelitian selanjutnya diharapkan dapat mengklasifikasikan data dalam jumlah besar dengan klasifikasi hotel unique type lebih dari 5 klasifikasi.

DAFTAR PUSTAKA

- Pariwisata Kini Jadi Andalan Pendukung Devisa Negara <http://www.kemenpar.go.id/asp/detil.asp?c=16&id=2959>
- E. Pitoska, "E-Tourism: Penggunaan Internet dan Teknologi Informasi dan Komunikasi dalam Pariwisata: Kasus Unit Hotel di Daerah Pinggiran," *Pariwisata di Eropa Tenggara*, vol. 2, hlm. 335–344, Desember 2013.
- E. Pantano dan LD Pietro, "Dari e-tourism ke ftourism: isu-isu yang muncul dari wisatawan negatif' ulasan online," *Journal of Hospitality and Tourism Technology*, vol. 4, tidak. 3, hlm. 211–227, 2013.
- D. Li Guo, D. Peng dan L. Ai Ping, "Algoritma Klasifikasi Teks Naive Bayes Baru," *TELKOMNIKA Jurnal Teknik Elektro Indonesia*, vol. 12, tidak. 2, hlm. 947-952, 1 Februari 2014.
- J. Samodra, S. Sumpeno dan M. Hariadi, "Klasifikasi dokumen teks berbahasa Indonesia dengan menggunakan naive bayes," dalam *Seminar Nasional Pendidikan Elektro, Informatika, dan IT*, 2009.
- AF Hidayatullah dan Azhari SN, "Analisis sentimen dan klasifikasi kategori terhadap tokoh publik di Twitter," dalam *Seminar Nasional Informatika 2014 (semnasIF 2014)*, Yogyakarta, 2014.
- V. Chandani, RS Wahono dan Purwanto, "Komparasi algoritma klasifikasi machine learning dan seleksi fitur pada analisis sentimen review film," *Jurnal Sistem Cerdas*, vol. 1, tidak. 1 Februari 2015.
- Risma Febriani, Devit Satria, Mustazihim "Rancang Bangun Game Cinta Indonesia Tentang Pakaian, Alat Musik dan Rumah Adat Indonesia Berbasis Android", *Jurnal Teknologi Komputer dan Informasi (JUTEKINF)*, Juni 2022.

Heliyanti Susana, Nana Suarna Fathurrohman, Kaslani, "Penerapan Model Klasifikasi Metode Naïve Bayes Terhadap Penggunaan Akses Internet", *JURISISTEKNI (Jurnal Sistem Informasi dan Teknologi Informasi)* Vol 4, No.1, January 2022: Hal 1– 8 ISSN.

Jiawei, H. (2012). *Data Mining: Concepts and Techniques*. Morgan Kaufmann.

Rizal Rachman, Rissa Nurfitriana Handayani, "Klasifikasi Algoritma Naive Bayes Dalam MemprediksiTingkat Kelancaran Pembayaran Sewa Teras UMKM", *JURNAL INFORMATIKA*, Vol.8 No.2 September 2021, Halaman 111~122.