

Penerapan Model Klasifikasi Untuk Prediksi Gender Berdasarkan Wajah Dengan Menggunakan Metode Naïve Bayes

Alvin Ancala Tirto Prayogo
Fakultas Teknik dan Ilmu Komputer
Universitas Kristen Krida Wacana
[Email: 412019006.alvin1@civitas.ukrida.ac.id](mailto:412019006.alvin1@civitas.ukrida.ac.id)

Penulis Korespondensi : Alvin Ancala Tirto Prayogo

Abstrak— Gender yang merupakan karakteristik dari seseorang adalah salah satu informasi yang banyak dibutuhkan pada berbagai bidang. Mudah bagi manusia untuk mengenali gender seseorang dengan melihat wajahnya, akan tetapi bagi computer tidak mudah untuk dapat mengenali gender seseorang tanpa adanya informasi tambahan yaitu ciri-ciri fisik dari orang tersebut. Penelitian ini membuat penerapan model klasifikasi untuk prediksi gender berdasarkan wajah dengan menggunakan metode Naïve Bayes pada RapidMiner.

Kata Kunci— Data Mining, Klasifikasi, ,Naïve Bayes, Gender.

Abstract— Gender which is a characteristic of a person is one of the much needed information in various fields. It is easy for humans to recognize a person's gender by looking at his face, but for computers it is not easy to recognize a person's gender without additional information, namely the physical characteristics of the person. This study makes the application of a classification model for gender prediction based on faces using the Naïve Bayes method on RapidMiner.

Keywords— Data Mining, Classification, Naïve Bayes, Gender



I. PENDAHULUAN

Pada zaman yang telah berkembang ini terutama pada bidang teknologi. Saat ini teknologi yang berkaitan dengan pengenalan wajah seseorang telah banyak dimanfaatkan pada aplikasi pengenalan data biologis, pencarian dan pengindeksan database citra dan video digital, keamanan ruangan dan lain lain (Asmara, Andjani, Rosiani, & Choirina, 2018). Data biologis dapat memberikan berbagai informasi dari seseorang dengan menggunakan informasi ciri-ciri dari fisik setiap orang yang berbeda seperti wajah. Wajah adalah salah satu pembeda yang dapat membuat setiap orang menjadi berbeda satu sama lain dan salah satu informasi yang dapat diambil dari wajah adalah gender.

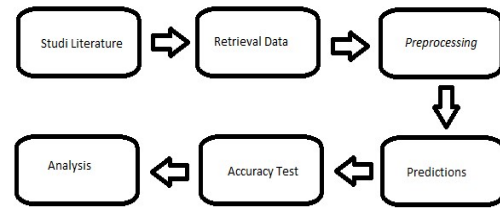
Gender adalah karakteristik dari setiap orang yang mencakup dengan jenis kelamin seseorang. Gender merupakan salah satu dari informasi yang banyak dibutuhkan dalam melakukan pengenalan seseorang karena informasi gender dibutuhkan pada berbagai bidang (Devito, Wihandika, & Widodo, 2019). Oleh karena itu telah banyak peneliti yang telah mengembangkan sebuah system klasifikasi gender untuk dapat mengetahui gender dari seseorang. Salah satu cara untuk dapat mengenali gender dari seseorang adalah dengan melihat wajah dari orang tersebut. Hal ini merupakan hal yang mudah untuk dilakukan bagi manusia, akan tetapi tidak mudah untuk dapat dilakukan oleh computer (Armandhani, Wihandika, & Rahman, 2019). Pada klasifikasi gender dapat dikenali gender seseorang adalah laki-laki atau perempuan berdasarkan informasi tambahan rambut dan bentuk wajah dari orang tersebut (Munarto & Darma, 2021).

Penulis tertarik untuk meneliti dalam menerapkan model klasifikasi untuk dapat memprediksikan gender seseorang berdasarkan ciri fisik wajah dari panjang rambut, panjang dan lebar kening, panjang dan lebar hidung, bibir, dan jarak hidung ke bibir dengan menggunakan metode Naïve Bayes pada aplikasi RapidMiner.

Metode naïve bayes adalah salah satu metode yang memiliki algoritma pembelajaran induktif yang paling efektif dan efisien untuk machine learning dan data mining (Syarli & Muin, 2016). Metode Naïve bayes adalah metode perhitungan probabilitas yang dilakukan secara statistic untuk dapat memprediksi sebuah peluang yang akan datang berdasarkan pengalaman atau permasalahan yang dihadapi sebelumnya (Susilo, Setiawan, Saputro, Purwadi, & Saifudin, 2021). Model Klasifikasi dengan metode naïve bayes sudah terbukti cocok dan banyak digunakan sebagai metode untuk memprediksi hasil dari suatu permasalahan (Asmara et al., 2018)(Syarli & Muin, 2016)(Susilo et al., 2021)(Heliyanti Susana, 2022).

II. Metode Penelitian

Tahapan penelitian yang dilakukan untuk menyelesaikan penelitian ini digambarkan pada Gambar 1 dibawah ini.



Gambar 1. Tahapan penelitian

Gambar 1 adalah rangkaian penelitian yang perlu dilakukan pada penelitian ini. Studi literature dilakukan untuk dapat memahami, informasi-informasi yang terkait dengan penelitian ini termasuk mencari penelitian terdahulu. Kemudian dilakukan pengumpulan dataset, dataset yang digunakan pada penelitian ini yaitu dengan menggunakan data sekunder yaitu data yang diambil secara tidak langsung oleh penulis melalui perantara website terbuka Kaggle. Pada tahap preprocessing dilakukan penyesuaian pada data untuk dapat digunakan pada aplikasi RapidMiner. Setelah data dapat digunakan, maka dilakukan prediksi dengan menggunakan klasifikasi metode Naïve Bayes. Tahap selanjutnya akan di buat akurasi dan hasilnya. Klasifikasi adalah proses untuk menemukan sebuah model yang dapat membedakan kelas data dengan tujuan untuk dapat memprediksi sebuah kelas objek yang labelnya tidak diketahui. Klasifikasi merupakan metode supervised learning, yaitu metode yang mencoba untuk menemukan hubungan antara atribut masukan dan atribut target (Hendrian, 2018).

Naïve Bayes adalah salah satu model klasifikasi. Naïve bayes merupakan pengklasifikasian yang dilakukan dengan menghitung probabilitas yang dilakukan secara statistic yang dikemukakan oleh seorang ilmuwan yang berasal dari Inggris yaitu Thomas Bayes untuk memprediksi sebuah peluang yang akan datang berdasarkan pengalaman atau permasalahan yang telah dihadapi sebelumnya. Metode ini memerlukan data sample yang digunakan sebagai data training untuk dapat membuat prediksi data selanjutnya. Keuntungan yang didapat dari penggunaan metode Naïve Bayes adalah metode ini hanya membutuhkan jumlah data training yang kecil untuk menemukan estimasi parameter yang diperlukan dalam proses pengklasifikasian (Nurmayanti, 2021).

RapidMiner adalah aplikasi perangkat lunak bersifat terbuka yang digunakan sebagai alat untuk melakukan data mining. Rapid miner dapat melakukan perhitungan metode Naïve Bayes secara otomatis dengan menggunakan operatornya saja.

III. Hasil dan Pembahasan

Telah didapatkan dataset ciri-ciri fisik wajah dari seseorang untuk menentukan gender dari orang tersebut. Dataset memiliki data sample sebanyak 5001 row.

Table 1. Dataset Ciri-ciri Wajah

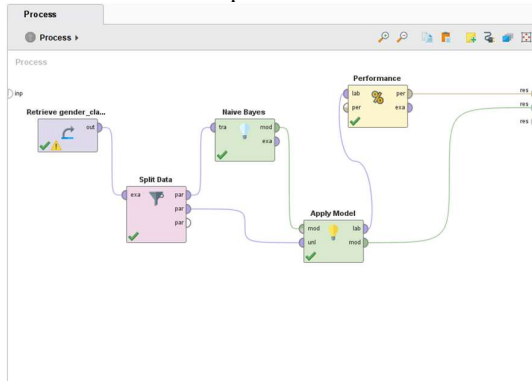
n	long hair	forehead width	forehead height	nose wide	nose long	lips thin	distance nose to lip long	gender
1	1	11.8	6.1	1	0	1	1	Male
2	0	14	5.4	0	0	1	0	Female
3	0	11.8	6.3	1	1	1	1	Male
4	0	14.4	6.1	0	1	1	1	Male
5	1	13.5	5.9	0	0	0	0	Female
6	1	13	6.8	1	1	1	1	Male
7	1	15.3	6.2	1	1	1	0	Male
8	0	13	5.2	0	0	0	0	Female
9	1	11.9	5.4	1	0	1	1	Female
10	1	12.1	5.4	0	0	0	0	Female
11	0	12.5	5.4	1	1	1	1	Male
12	1	15.5	5.8	1	1	1	1	Male
13	0	14.7	5.2	1	1	1	1	Male
14	1	14.5	6.7	0	1	1	1	Male
15	1	14.2	6.5	0	0	0	0	Female
16	0	12.5	5.2	1	1	1	1	Male
17	1	15.2	6	1	1	1	1	Male
18	1	14	5.9	0	0	0	0	Female
19	1	11.9	5.2	0	0	0	0	Female
20	1	14.4	5.6	1	0	1	1	Male
21	1	14.6	6.3	1	1	1	1	Male
22	1	12.8	7	1	1	1	1	Male
23	1	12.9	6.4	1	0	0	1	Female
24	0	12.2	6	1	1	0	1	Male
25	0	14.2	5.3	0	0	0	0	Female



26	1	12.3	6.3	0	0	0	0	Female
27	1	14.2	6.2	0	0	0	0	Female
28	1	12.5	5.2	0	0	0	0	Female
29	1	15.5	5.7	1	1	1	0	Male
30	1	15.3	6.3	1	1	1	1	Male
...



Saat data dimasukkan ke aplikasi RapidMiner, role pada kolom gender akan dijadikan sebagai label sebagai kolom yang akan diprediksi. Berikut desain dari RapidMiner



Gambar 2. Desain Data Mining

Berdasarkan data diatas dijelaskan alur dari model klasifikasi dengan Naive Bayes yang telah dibuat. Untuk memudahkan Proses analisa, akan digunakan operator split data untuk memecah dataset menjadi 2 bagian yaitu 80% dari dataset akan dijadikan sebagai data training dan 20% akan dijadikan sebagai data testing.

The screenshot shows the Performance Vector window in RapidMiner. It displays the following table:

Criterion	Value
accuracy	98.00%
precision	
recall	
AUC (symmetric)	
AUC	
AUC (asymmetric)	

	true Male	true Female	class precision
pred Male	490	10	98.00%
pred Female	10	490	98.00%
class recall	98.00%	98.00%	

Gambar 3. Hasil Akurasi

Berdarkan gambar diatas dapat disimpulkan bahwa tingkat akurasi adalah 98% dengan rincian detail seperti berikut :

- Hasil prediksi laki-laki dan ternyata laki-laki sebanyak 490
- Hasil prediksi laki-laki dan ternyata bukan laki-laki sebanyak 10
- Hasil prediksi perempuan dan ternyata perempuan sebanyak 490
- Hasil prediksi laki-laki dan ternyata bukan perempuan sebanyak 10

IV. Kesimpulan

Dari penelitian yang sudah dilakukan dapat ditarik kesimpulan bahwa penerapan model klasifikasi metode Naive Bayes pertama dengan mencari dan memasukan dataset serta melakukan preprocessing pada dataset terlebih dahulu pada saat dimasukkan ke RapidMiner. Kemudian dilanjutkan dengan penambahan operator seperti split data untuk memisah data menjadi beberapa bagian, Naive Bayes untuk metode klasifikasinya, apply model untuk testingnya, dan performance untuk menunjukkan akurasi.

Akurasi yang didapat pada penelitian dengan menggunakan model klasifikasi metode Naive Bayes adalah 98%. Hasil prediksi laki-laki dan ternyata laki-laki sebanyak 490. Hasil prediksi laki-laki dan ternyata bukan laki-laki sebanyak 10. Hasil prediksi perempuan dan ternyata perempuan sebanyak 490. Hasil prediksi laki-laki dan ternyata bukan perempuan sebanyak 10.

Daftar Pustaka

- Aprilia Triase; Sriani, Sriani, Rima; Triase. 2017. "Penentuan Tempat Menginap Dengan Menggunakan Fuzzy Multiple Attribute Decision Making." *Algoritma: Jurnal Ilmu Komputer Dan Informatika* 1, no. 1: 30–34. <http://jurnal.uinsu.ac.id/index.php/algoritma/article/view/1306>.
- Borman, Rohmat Indra, and Helmi Fauzi. 2018. "Dalam Sistem Pendukung Keputusan Penerima Beasiswa." *CESS: Journal of Computer Engineering, System and Science* 3, no. 1: 17–22.
- Fauzan, Reza, Yoenie Indrasary, and Nonik Muthia. 2018. "Sistem Pendukung Keputusan Penerimaan Beasiswa Bidik Misi Di POLIBAN Dengan Metode SAW Berbasis Web." *Jurnal Online Informatika* 2, no. 2: 79. <https://doi.org/10.15575/join.v2i2.101>.
- Fitriana, Julia, Eva Faja Ripanti, and Tursina Tursina. 2018. "Sistem Pendukung Keputusan Pemilihan Mahasiswa Berprestasi Dengan Metode Profile Matching." *Jurnal Sistem Dan Teknologi Informasi (JUSTIN)* 6, no. 4: 153. <https://doi.org/10.26418/justin.v6i4.27113>.

- Ramadhon, Raka, Riswan Jaenudin, and Siti Fatimah. 2017. "Pengaruh Beasiswa Terhadap Motivasi Belajar Mahasiswa Pendidikan Ekonomi Universitas Sriwijaya." *Jurnal Profit* 4, no. 2: 203–13.
<https://media.neliti.com/media/publications/205759-pengaruh-beasiswa-dan-motivasi-belajar-t.pdf>.
- Rinaldhi, Galih Eka. 2011. "Penerapan Metode Simple Additive Weighting (SAW) Untuk Sistem Pendukung Keputusan Penentuan Penerimaan Beasiswa Bantuan Siswa Miskin (Bsm) Pada Sma Negeri 1 Subah Kab . Batang." *Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Dian Nuswantoro Semarang*, 1–9.
- Tasril, Virdyra. 2018. "Sistem Pendukung Keputusan Pemilihan Penerimaan Beasiswa Berprestasi Menggunakan Metode Elimination Et Choix Traduisant La Realite." *INTECOMS: Journal of Information Technology and Computer Science* 1, no. 1: 100–109.
<https://doi.org/10.31539/intecom.s.v1i1.163>.
- Veza, Okta, and Nofri Yudi Arifin. 2020. "Sistem Pendukung Keputusan Calon Mahasiswa Non Aktif Dengan Metode Simple Additive Weighting." *Jurnal Industri Kreatif (JIK)* 3, no. 02: 71–78.
<https://doi.org/10.36352/jik.v3i02.29>.
- Wang, Vioni Vevila, Anggi Srimurdianti Sukanto, and Enda Esyudha Pratama. 2019. "Sistem Pendukung Keputusan Seleksi Mahasiswa Penerima Beasiswa BBP-PPA Dengan Metode TOPSIS Pada Fakultas Teknik UNTAN." *Jurnal Sistem Dan Teknologi Informasi (JUSTIN)* 7, no. 2: 105.
<https://doi.org/10.26418/justin.v7i2.29656>.