



Klasifikasi Gender Berdasarkan *Fingerprint* Menggunakan Metode *Naive Bayes Classifier*

Gender Classification Based on Fingerprint Using Naive Bayes Classifier Method

Cindhy Herumawan* dan Efi Anisa

Program Studi Teknik Informatika, Sekolah Tinggi Teknologi Duta Bangsa, Jl. Kalibaru Timur, Kalibaru Medan Satria, Jawa Barat 17530, Indonesia

Informasi artikel:

Diterima:
28/09/2022
Direvisi:
21/11/2022
Disetujui:
23/11/2022

Abstract

No two fingerprints are identical, as everyone has their own characteristics. The most fundamental problem lies in the results of the fingerprint image, typically due to inconsistencies in the emphasis of the fingerprint and the position of the fingerprint, resulting in inconsistencies in the thickness of the black line and shifting positions, which negatively impact the overall performance of the system. To solve this issue, research is required on a classifier that assumes all attributes exist independently. The NBC (*Naive Bayes Classifier*) is a classifier based on the assumption that all attributes are independent. The NBC method for gender classification based on fingerprints consists of three steps. The initial step is to evaluate the quality of the image to be processed. This is demonstrated by the consistency of the grayscale values, which are not skewed when converted to a binary image. The second is the selection of data that exhibits no data deviation, which also leads to errors in the classification procedure that follows. With the existence of machine learning, class-based measurement formulations can be acquired through training. Even with unbalanced data, it is preferable to use NBC for classification purposes.

Keywords: *Naive Bayes Classifier*, normal distribution, deviation.

SDGs:



Abstrak

Fingerprint adalah hal yang unik dimana setiap orang mempunyai ciri tersendiri dan tidak ada yang sama. Permasalahan yang paling mendasar adalah terletak pada hasil citra sidik jarinya biasanya hal ini karena tidak konsisten pada penekanan sidik jari dan posisi sidik jari sehingga mengakibatkan ketebalan warna garis hitam yang tidak konsisten dan posisi yang berubah ubah, sehingga pada akhirnya mempengaruhi kinerja sistem secara keseluruhan. Untuk mengatasi masalah ini, maka dibutuhkan penelitian tentang *classifier* yang mengasumsikan semua atribut berdiri sendiri. NBC (*Naive Bayes Classifier*) adalah suatu *classifier* yang mengasumsikan semua atribut berdiri sendiri. Penerapan metode NBC berbasis *fingerprint* untuk klasifikasi gender di bagi menjadi 3 proses. Pertama adalah memastikan apakah citra yang akan di proses adalah citra yang berkualitas baik. Hal ini dapat dilihat dengan konsistennya nilai *grayscale*, sehingga tidak menjadi bias ketika akan dirubah menjadi citra biner. Kedua adalah pemilihan data yang tidak mengalami deviasi atau penyimpangan data, hal ini juga mengakibatkan kesalahan dalam proses klasifikasi selanjutnya. Ketiga dengan adanya *machine learning* akan didapat rumusan untuk pengukuran dari kelas melalui *training*. Sehingga dengan menggunakan NBC menjadi lebih baik dalam proses klasifikasi walau dengan data *imbalance*.

Kata Kunci: *Naive Bayes Classifier*, distribusi normal, deviasi.

*Penulis Korespondensi. Tel: +6221 2908 2747 ; Hp: +62 881 1696 206
email : cindhyheru@sttdb.ac.id



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)

4. PENDAHULUAN

Saat Sidik jari setiap orang adalah unik dan tidak ada yang sama, dan cenderung untuk tidak berubah dalam pertumbuhan manusia. Setiap orang memiliki struktur sidik jari yang unik (Kusumawati, D'arofah dan Pramana, 2019). Sidik jari adalah suatu bukti yang tidak diragukan lagi dan yang paling dapat diandalkan dan tidak berubah dari lahir sampai meninggal (Gungadin, 2007). Karena sifat itulah maka sidik jari dapat digunakan sebagai bukti dalam persidangan di seluruh dunia. Sidik jari adalah salah satu teknologi biometrik yang paling matang dan dianggap bukti yang sah bukti di pengadilan di seluruh dunia (Gornale, Basavanna dan Kruthi, 2017). Dalam bidang *forensic* dapat sebagai alat bantu untuk mengurangi jumlah para tersangka, dan di bidang keamanan sidik jari digunakan sebagai identitas dari suatu individu. Salah satu pengembangan dari penelitian sidik jari adalah menganalisa sidik jari untuk menentukan gender.

Gondara dan Kumar (Godara dan Kumar, 2019) telah melakukan penelitian sidik jari berdasarkan gender dengan menganalisa *Features extracted* : Kepadatan lembah, lebar bukit, lebar lembah dengan ditambahkan *classifier* : SVM (*Support Vector Machine*). Metode lain untuk menganalisa citra sidik jari adalah dengan citra transformasi, seperti DWT (*Discrete Wavelet Transform*) yang sudah di gunakan oleh, dan dikembangkan juga oleh peneliti yang lain dengan kombinasi pada *classifiernya* (Al Rivan, Rachmat dan Ayustin, 2020). Metode analisa citra yang lain seperti FFT (*Fast Fourier Transform*), DCT (*Discrete Cosin Transform*) pada Kaur, serta metode kombinasi lainnya (Syahputra, Yanris dan Irmayani, 2022).

Metode *Features extracted* : Kepadatan lembah, lebar bukit, lebar lembah dengan ditambahkan *classifier* : SVM (*Support Vector Machine*) yang telah dilakukan dan sudah menghasilkan tingkat akurasi hingga 91% (Gornale, Basavanna dan Kruthi, 2017). Kelebihan dari klasifikasi SVM adalah *non linier classifier* dan sering digunakan untuk memproduksi klasifikasi yang sangat bagus (Syaputri, Irwandi dan Mustakim, 2020). Sedangkan kelemahannya

adalah untuk data yang *imbalance* untuk data minoritas akurasi akan menurun (Handayani dan Pribadi, 2015).

Untuk mengatasi masalah ini maka dibutuhkan *classifier* yang mengasumsikan semua atribut berdiri sendiri. NBC (*Naive Bayes Classifier*) adalah suatu *classifier* yang mengasumsikan semua atribut berdiri sendiri (Windarti dan Suradi, 2019; Siregar, Siregar dan Sudirman, 2020). Dalam penelitian yang dilakukan menggunakan NBC sebagai pengklasifikasiannya yang bertujuan untuk meningkatkan akurasi dari analisa sidik jari.

5. METODOLOGI

2.1. Analisa kebutuhan

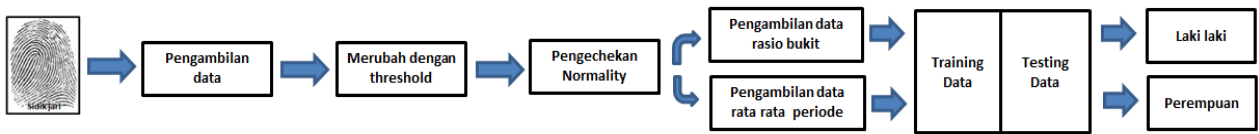
Dalam perancangan penelitian akan dibahas masing masing proses untuk menggambarkan lebih detail prosesnya seperti pada Gambar 1. Untuk dapat melakukan penelitian seperti pada Gambar 1, perlu adanya persiapan data atau *software*. Hal yang perlu disiapkan adalah:

- 1) Dataset.
- 2) Dataset diambil dari NISTS (*National Institute for the Study of Transfer Students*). Data yang diambil hanya data dari ibu jari untuk tangan sebelah kanan dan jumlah data sekitar 150 data sidik jari.
- 3) Untuk pengambilan data panjang bukit dan lembah menggunakan *software Image-J*.
- 4) Pengolahan data citra dan perhitungan probabilitas *Naive Bayes* menggunakan microsoft excel.
- 5) Untuk statistik analisa menggunakan Minitab versi 15.
- 6) Sebuah Laptop dengan minimal Pentium dual core.

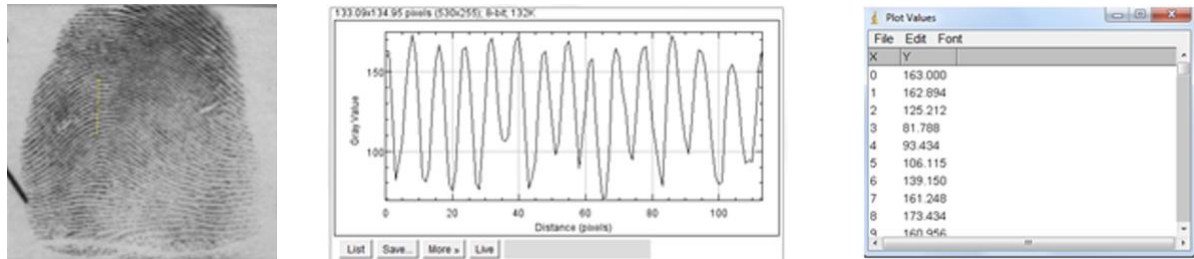
2.2. Perancangan Penelitian

2.2.1. Teknik pengambilan Data

Pengambilan data dilakukan dengan menarik garis pada citra sidik jari (Gambar 2a), maka akan didapat grafik nilai *grayscale* pada tiap-tiap *pixel* garis (Gambar 2b). Citra yang bagus memiliki amplitudo cenderung datar dan tidak ditemukan riak pada ujung gelombang. Gambar 2c adalah nilai *grayscale* (*column Y*) dari tiap-tiap *pixel* sepanjang garis tersebut (*column X*).



Gambar 1. Gambaran model penelitian



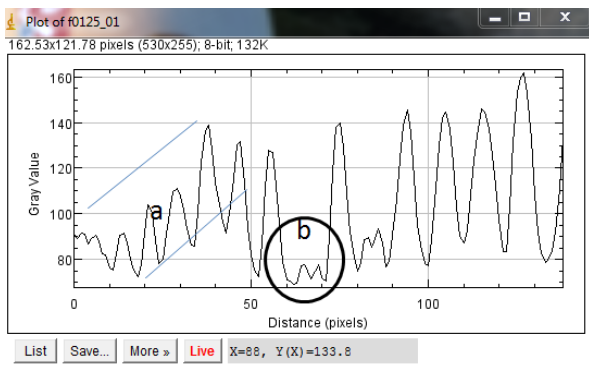
(a). Sample sidik jari

(b). Grafik nilai grayscale terhadap pixel

(c). Nilai grayscale (Y) pada tiap nomer urutan pixel (X)

Gambar 2. Pengambilan data

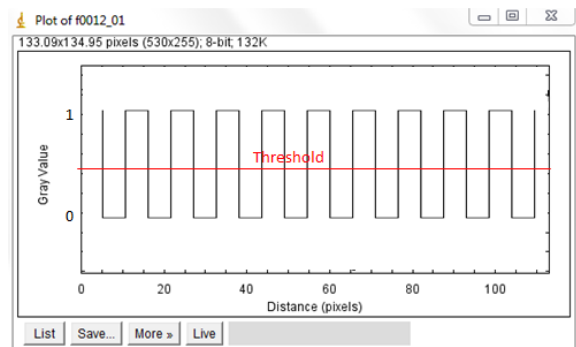
Pada Gambar 3 menampilkan hasil citra dari software Image J yang buruk atau data tidak konsisten terhadap nilai grayscale, dimana “a” adalah amplitudo cenderung naik atau tidak datar, “b” adalah adanya riak pada ujung gelombang, sehingga data tidak dapat dihitung batasan hitam dan putihnya. Apabila pada citra tidak didapat grafik yang baik maka citra dinyatakan tidak layak.



Gambar 3. Grafik grayscale dari citra yang buruk

2.2.2. Data Biner dengan Threshold

Untuk menentukan nilai threshold adalah dengan menghitung jumlah nilai dari semua data tersebut ($\sum Y$) dibagi banyaknya data (n) (Zain dan Sibaroni, 2019). Setelah mendapatkan nilai threshold, data dirubah dari nilai grayscale menjadi nilai 1 dan 0, 1 mewakili nilai putih dan 0 nilai hitam Gambar 4. Nilai 0 apabila nilai grayscale < threshold, dan 1 untuk sebaliknya.



Gambar 4. Grafik citra biner dengan threshold

2.2.3. Pengecekan Normality

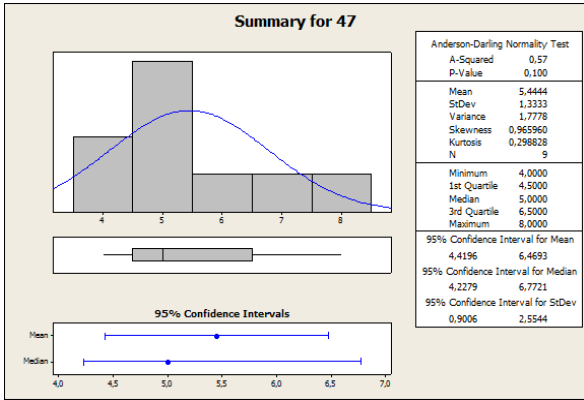
Untuk memastikan data yang kita dapat tidak mempunyai penyimpangan maka data tersebut kita lakukan test normality, menggunakan minitab untuk melihat nilai P (P value), apabila P value > 0,05; maka data tersebut normal atau tidak ada penyimpangan data yang besar, bila P value < 0,05 berarti data mempunyai penyimpangan atau tidak normal distribusi. Apabila data tersebut tidak normal, maka perlu perbaikan data ulang atau data di abaikan. Hasil pengukuran dengan software Minitab dapat dilihat pada Gambar 5.

2.2.4. Pengambilan Data

Berdasarkan Gambar 6, dimana B = panjang bukit (pixel), L = panjang lembah (pixel), T = panjang periode (pixel), dan P = panjang data (pixel). Sehingga Rasio bukit adalah:

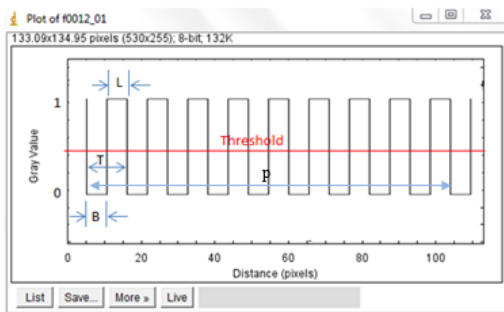
$$\text{Rasio bukit} = \frac{B}{T} \quad (1)$$

$$\text{Rata - rata panjang periode} = \frac{B}{\Sigma T} \quad (2)$$



	C1	C2	C3	C4
	no citra	12	47	50
4	4	4	5	6
5	5	4	5	8
6	6	4	6	5
7	7	4	4	5
8	8	3	5	6
9	9	4	4	7
10	10	4		7
11	11	4		7
12	12	5		5
13	13	4		5
14	14			5
15	15			10
16	16			9
17	17			5
18	18			4
19	19			4

Gambar 5. Hasil pengukuran normality dengan Minitab



Gambar 6. Grafik pengukuran

2.2.5. NBC (Naive Bayes Classifier) Analisa

NBC digunakan untuk mendapatka Y (output klasifikasi) dari variable X. Output klasifikasi adalah gender laki laki atau perempuan. Sedangkan variable X adalah nilai rata rata rasio bukit dan rata rata panjang periode. Pertama variabel X data harus dirubah menjadi data nilai rata rata / mean (μ) dan standar deviasi (σ) dari data training.

3. HASIL DAN PEMBAHASAN

3.1. Kelas Perempuan / Female

Setiap masing masing citra akan di ukur jarak dari masing masing bukit dan lembah satu sampel garis pengukuran. Dari hasil ukur tersebut didapat analisa grafiknya serta data tablenya seperti pada Tabel 1. dengan dataset NISTS kelas perempuan hanya mempunyai 32 record.

Berdasarkan Gambar 7 terdapat sebuah citra sidik jari kemudian di tarik satu garis sembarang untuk mengukur berapa panjang bukit dan lembah dalam bentuk grafik. Dari grafik didapat nilai grayscale terhadap jarak (pixel). Data tersebut bisa juga didapat dalam bentuk tabel yang akan diproses untuk pengukurannya. Proses ini menggunakan aplikasi Image-J.

Data rasio bukit dan panjang periode setelah melalui normality test dihasilkan tabel data yang akan dijadikan data training (lihat Tabel 2).

3.2. Kelas Laki laki / Male

Dataset NISTS kelas laki-laki disajikan dalam Tabel 3 yang diambil sebanyak 50 record. Sedangkan data rasio bukit dan panjang periode setelah melalui normality test diperlihatkan pada Tabel 4.

3.3. NBC (Naive Bayes Classifier) Analisa dan Akurasi

Didalam NBC untuk menilai hasil dari machine learning ini akan dilakukan dengan dua tahap. Yang pertama adalah training data, hal ini digunakan untuk menentukan rumusan dalam penentuan output kelas dari masing masing data (X). Untuk penelitian ini di gunakan semua data sebagai training data seperti pada Tabel 5. Dari Tabel 5, maka prior masing masing kelas adalah:

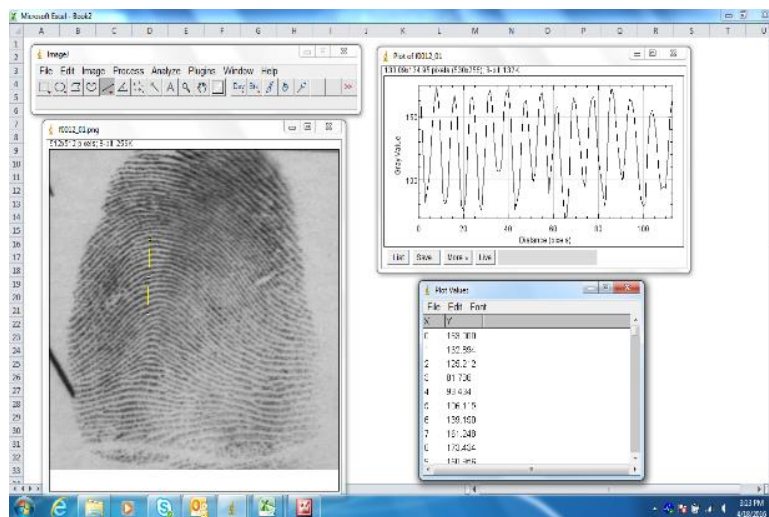
Untuk Prior Female: $P(C_F) = \frac{\Sigma C_F}{n} = \frac{11}{32} = 0,34$

Untuk Prior Male : $P(C_M) = \frac{\Sigma C_M}{n} = \frac{21}{32} = 0,66$

Dari Tabel 6 adalah sebagai estimasi parameter dari penghitungan Likelihood dari tiap tiap X, yaitu nilai rasio dan periode dengan menggunakan distribusi Gaussian.

Tabel 1. Dataset fingerprint female

No	No. of Image	Gender	No	No. of Image	Gender
1	f0012_01	F	17	f0847_01	F
2	f0136_01	F	18	f0883_01	F
3	f0141_01	F	19	f0890_01	F
4	f0193_01	F	20	f0895_01	F
5	f0215_01	F	21	f0932_01	F
6	f0240_01	F	22	f0979_01	F
7	f0293_01	F	23	f1047_01	F
8	f0302_01	F	24	f1125_01	F
9	f0391_01	F	25	f1162_01	F
10	f0413_01	F	26	f1195_01	F
11	f0471_01	F	27	f1244_01	F
12	f0609_01	F	28	f1272_01	F
13	f0615_01	F	29	f1328_01	F
14	f0647_01	F	30	f1562_01	F
15	f0735_01	F	31	f1618_01	F
16	f0790_01	F	32	f1644_01	F



Gambar 7. Sampel pengukuran sidik jari

Tabel 2. Rasio dan periode female

No	No. of Image	Threshold	Rasio (%)	Period (pixel)	Test Normality Rasio	Test Normality period	Result
136	f0136_01	107.07	49.20	9.13	0.482	0.11	Pass
215	f0215_01	99.72	46.25	8.50	0.222	0.489	Pass
240	f0240_01	122.51	49.95	14.50	0.054	0.973	Pass
302	f0302_01	130.24	48.32	11.25	0.268	0.259	Pass
735	f0735_01	141.66	42.37	10.22	0.27	0.167	Pass
790	f0790_01	107.77	47.58	9.43	0.182	0.323	Pass
895	f0895_01	97.13	49.89	13.38	0.257	0.487	Pass
932	f0932_01	110.95	50.52	15.64	0.169	0.271	Pass
979	f0979_01	80.20	48.59	12.40	0.236	0.068	Pass
1244	f1244_01	104.08	48.06	11.26	0.456	0.094	Pass
1328	f1328_01	96.16	50.58	9.78	0.614	0.578	Pass

Tabel 3. Dataset Fingerprint Male

No.	No. of Image	Gender	No.	No. of Image	Gender
1	f0001_01	M	26	f0421_01	M
2	f0010_01	M	27	f0445_01	M
3	f0015_01	M	28	f0446_01	M
4	f0047_01	M	29	f0447_01	M
5	f0050_01	M	30	f0467_01	M
6	f0068_01	M	31	f0470_01	M
7	f0113_01	M	32	f0473_01	M
8	f0125_01	M	33	f0490_01	M
9	f0137_01	M	34	f0493_01	M
10	f0180_01	M	35	f0505_01	M
11	f0223_01	M	36	f0604_01	M
12	f0235_01	M	37	f0611_01	M
13	f0252_01	M	38	f0617_01	M
14	f0261_01	M	39	f0650_01	M
15	f0264_01	M	40	f0675_01	M
16	f0269_01	M	41	f0687_01	M
17	f0270_01	M	42	f0723_01	M
18	f0288_01	M	43	f0764_01	M
19	f0289_01	M	44	f0766_01	M
20	f0376_01	M	45	f0800_01	M
21	f0381_01	M	46	f0802_01	M
22	f0382_01	M	47	f0805_01	M
23	f0402_01	M	48	f0814_01	M
24	f0408_01	M	49	f0821_01	M
25	f0414_01	M	50	f0824_01	M

Tabel 4. Rasio dan periode male

No	No. of Image	Threshold	Rasio (%)	Period (pixel)	Test Normality Rasio	Test Normality period	Result
1	f0001_01	116.51	55.96	13.38	0.173	0.55	Pass
15	f0015_01	107.77	54.64	11.25	0.686	0.236	Pass
47	f0047_01	72.87	53.91	11.78	0.84	0.216	Pass
125	f0125_01	107.23	51.60	10.11	0.596	0.362	Pass
137	f0137_01	62.38	61.85	9.89	0.666	0.201	Pass
223	f0223_01	74.20	59.58	10.92	0.516	0.095	Pass
252	f0252_01	79.22	52.20	11.18	0.686	0.292	Pass
270	f0270_01	85.42	50.83	10.71	0.064	0.445	Pass
376	f0376_01	106.21	57.42	11.57	0.413	0.09	Pass
382	f0382_01	142.95	55.06	10.00	0.722	0.218	Pass
402	f0402_01	127.46	52.09	13.36	0.859	0.125	Pass
473	f0473_01	87.07	54.91	9.89	0.206	0.716	Pass
604	f0604_01	133.22	51.54	11.63	0.371	0.161	Pass
650	f0650_01	84.01	53.22	9.08	0.22	0.104	Pass
675	f0675_01	120.94	52.15	11.38	0.565	0.731	Pass
687	f0687_01	104.62	51.90	12.33	0.567	0.421	Pass
723	f0723_01	126.95	55.83	10.00	0.792	0.665	Pass
764	f0764_01	119.50	55.15	12.20	0.906	0.536	Pass
802	f0802_01	119.85	53.86	10.22	0.481	0.48	Pass
814	f0814_01	66.50	51.51	9.31	0.063	0.416	Pass
821	f0821_01	104.70	52.09	11.71	0.551	0.743	Pass

Tabel 5. Data training

No	No. of Image	Threshold	Rasio (%)	Period (pixel)	Test Normality Rasio	Test Normality period	Result	Gender
1	f0001_01	116.51	55.96	13.38	0.173	0.55	Pass	M
15	f0015_01	107.77	54.64	11.25	0.686	0.236	Pass	M
47	f0047_01	72.87	53.91	11.78	0.84	0.216	Pass	M
125	f0125_01	107.23	51.60	10.11	0.596	0.362	Pass	M
137	f0137_01	62.38	61.85	9.89	0.666	0.201	Pass	M
223	f0223_01	74.20	59.58	10.92	0.516	0.095	Pass	M
252	f0252_01	79.22	52.20	11.18	0.686	0.292	Pass	M
270	f0270_01	85.42	50.83	10.71	0.064	0.445	Pass	M
376	f0376_01	106.21	57.42	11.57	0.413	0.09	Pass	M
382	f0382_01	142.95	55.06	10.00	0.722	0.218	Pass	M
402	f0402_01	127.46	52.09	13.36	0.859	0.125	Pass	M
473	f0473_01	87.07	54.91	9.89	0.206	0.716	Pass	M
604	f0604_01	133.22	51.54	11.63	0.371	0.161	Pass	M
650	f0650_01	84.01	53.22	9.08	0.22	0.104	Pass	M
675	f0675_01	120.94	52.15	11.38	0.565	0.731	Pass	M
687	f0687_01	104.62	51.90	12.33	0.567	0.421	Pass	M
723	f0723_01	126.95	55.83	10.00	0.792	0.665	Pass	M
764	f0764_01	119.50	55.15	12.20	0.906	0.536	Pass	M
802	f0802_01	119.85	53.86	10.22	0.481	0.48	Pass	M
814	f0814_01	66.50	51.51	9.31	0.063	0.416	Pass	M
821	f0821_01	104.70	52.09	11.71	0.551	0.743	Pass	M
136	f0136_01	107.07	49.20	9.13	0.482	0.11	Pass	F
215	f0215_01	99.72	46.25	8.50	0.222	0.489	Pass	F
240	f0240_01	122.51	49.95	14.50	0.054	0.973	Pass	F
302	f0302_01	130.24	48.32	11.25	0.268	0.259	Pass	F
735	f0735_01	141.66	42.37	10.22	0.27	0.167	Pass	F
790	f0790_01	107.77	47.58	9.43	0.182	0.323	Pass	F
895	f0895_01	97.13	49.89	13.38	0.257	0.487	Pass	F
932	f0932_01	110.95	50.52	15.64	0.169	0.271	Pass	F
979	f0979_01	80.20	48.59	12.40	0.236	0.068	Pass	F
1244	f1244_01	104.08	48.06	11.26	0.456	0.094	Pass	F
1328	f1328_01	96.16	50.58	9.78	0.614	0.578	Pass	F

Tabel 6. Mean dan Standar Deviation

Gender	Rasio		Periode	
	Mean	Standar Deviation	Mean	Standar Deviation
Male	54.16	2.85	11.04	1.20
Female	48.30	2.37	11.41	2.32

3.4. Testing Data

Data training di uji dengan perhitungan posterior masing masing kelas, maka nilai tertinggi dari posterior yang akan dianggap sebagai output dari kelas (lihat Tabel 7). Keterangan pada Tabel 7:

- Prior M : Prior male

- Prior F : Prior female
- Likelihood ratio M : probabilitas dari X =ratio bukit dari kelas male
- Likelihood ratio F : probabilitas dari X =ratio bukit dari kelas female
- Likelihood period M : probabilitas dari X =panjang periode dari kelas male

Tabel 7. Hasil testing

No. of Image	Gender	Rasio (%)	Period (pixel)	Prior F	Prior M	Likelihood Ratio-F	Likelihood Period-F	Posterior F	Likelihood Ratio-M	Likelihood Period-M	Posterior M	Result
f0001_01	M	55.96	13.38	0.34	0.66	0.0009	0.1200	3.71E-05	0.1148	0.0506	0.003834	Male
f0015_01	M	54.64	11.25	0.34	0.66	0.0047	0.1712	0.000272	0.1382	0.3269	0.02981	Male
f0047_01	M	53.91	11.78	0.34	0.66	0.0102	0.1695	0.000589	0.1397	0.2753	0.02538	Male
f0125_01	M	51.60	10.11	0.34	0.66	0.0639	0.1468	0.003188	0.0937	0.2457	0.015187	Male
f0137_01	M	61.85	9.89	0.34	0.66	0.0000	0.1388	6.33E-10	0.0036	0.2103	0.000504	Male
f0223_01	M	59.58	10.92	0.34	0.66	0.0000	0.1679	1.15E-07	0.0228	0.3301	0.004965	Male
f0252_01	M	52.20	11.18	0.34	0.66	0.0436	0.1708	0.00253	0.1106	0.3296	0.024057	Male
f0270_01	M	50.83	10.71	0.34	0.66	0.0951	0.1641	0.005308	0.0709	0.3196	0.014949	Male
f0376_01	M	57.42	11.57	0.34	0.66	0.0001	0.1712	5.92E-06	0.0725	0.3012	0.014422	Male
f0382_01	M	55.06	10.00	0.34	0.66	0.0029	0.1428	0.000139	0.1333	0.2277	0.020031	Male
f0402_01	M	52.09	13.36	0.34	0.66	0.0468	0.1208	0.001923	0.1077	0.0521	0.003704	Male
f0473_01	M	54.91	9.89	0.34	0.66	0.0035	0.1386	0.000163	0.1354	0.2093	0.01871	Male
f0604_01	M	51.54	11.63	0.34	0.66	0.0664	0.1709	0.003855	0.0917	0.2951	0.017861	Male
f0650_01	M	53.22	9.08	0.34	0.66	0.0195	0.1040	0.000691	0.1328	0.0879	0.007705	Male
f0675_01	M	52.15	11.38	0.34	0.66	0.0452	0.1716	0.002634	0.1092	0.3194	0.02302	Male
f0687_01	M	51.90	12.33	0.34	0.66	0.0533	0.1586	0.002872	0.1023	0.1866	0.012591	Male
f0723_01	M	55.83	10.00	0.34	0.66	0.0011	0.1428	5.26E-05	0.1180	0.2277	0.017732	Male
f0764_01	M	55.15	12.20	0.34	0.66	0.0026	0.1620	0.000142	0.1319	0.2089	0.018186	Male
f0802_01	M	53.86	10.22	0.34	0.66	0.0107	0.1506	0.000548	0.1395	0.2628	0.024191	Male
f0814_01	M	51.51	9.31	0.34	0.66	0.0674	0.1141	0.002615	0.0909	0.1171	0.007023	Male
f0821_01	M	52.09	11.71	0.34	0.66	0.0468	0.1701	0.002706	0.1078	0.2839	0.020198	Male
f0136_01	F	49.20	9.13	0.34	0.66	0.1568	0.1063	0.005665	0.0307	0.0940	0.001904	Female
f0215_01	F	46.25	8.50	0.34	0.66	0.1159	0.0784	0.003091	0.0030	0.0354	6.93E-05	Female
f0240_01	F	49.95	14.50	0.34	0.66	0.1321	0.0709	0.003183	0.0471	0.0053	0.000165	Female
f0302_01	F	48.32	11.25	0.34	0.66	0.1684	0.1712	0.009801	0.0171	0.3269	0.003682	Female
f0735_01	F	42.37	10.22	0.34	0.66	0.0074	0.1506	0.000377	0.0000	0.2628	4.59E-06	Female
f0790_01	F	47.58	9.43	0.34	0.66	0.1607	0.1194	0.006524	0.0097	0.1347	0.000861	Female
f0895_01	F	49.89	13.38	0.34	0.66	0.1345	0.1196	0.005468	0.0456	0.0498	0.001498	Female
f0932_01	F	50.52	15.64	0.34	0.66	0.1088	0.0327	0.001209	0.0618	0.0002	9E-06	Female
f0979_01	F	48.59	12.40	0.34	0.66	0.1671	0.1567	0.008904	0.0207	0.1755	0.0024	Female
f1244_01	F	48.06	11.26	0.34	0.66	0.1675	0.1713	0.009756	0.0142	0.3264	0.003051	Female
f1328_01	F	50.58	9.78	0.34	0.66	0.1061	0.1342	0.004841	0.0636	0.1907	0.008	Male

- Likelihood period F : probabilitas dari X =panjang periode dari kelas female
- Posterior M : Nilai kemungkinan dari kelas male
- Posterior F : Nilai kemungkinan dari kelas female
- Result adalah Hasil akhir, diambil dari nilai kemungkinan yang terbesar dari masing masing kelas.

Dari hasil testing di temukan satu record yang hasilnya tidak sesuai. Sehingga Akurasi rasio adalah:

$$Akurasi Rasio = \frac{31}{32} \times 100\% = 96,87\%$$

Atau bisa dikatakan dari 32 data yang di test, 31 data benar dan 1 salah.

4. SIMPULAN

Penerapan metode NBC berbasis *fingerprint* untuk klasifikasi gender di bagi menjadi 3 proses. Pertama adalah memastikan apakah citra yang akan di proses adalah citra yang berkualitas baik. Hal ini dapat dilihat dengan kosistennya nilai *grayscale*, sehingga tidak menjadi bias ketika akan dirubah menjadi citra biner. Kedua adalah pemilihan data yang tidak mengalami deviasi atau penyimpangan data, hal ini juga mengakibatkan kesalahan dalam proses klasifikasi selanjutnya. Ketiga dengan adanya *machine learning* akan

didapat rumusan untuk pengukuran dari kelas melalui *training*.

Dari tiga proses ini didapat kesimpulan pemilihan citra dengan visual digunakan untuk mencegah citra yang buruk yang memungkinkan kegagalan dalam penelitian. Mencegah deviasi data dengan *normality* test untuk memastikan data dalam distribusi normal atau tidak. Dan dengan menggunakan NBC menjadi lebih baik dalam proses klasifikasi walau dengan data *imbalance*.

DAFTAR PUSTAKA

- Al Rivan, M.E., Rachmat, N. dan Ayustin, M.R. (2020) 'Klasifikasi Jenis Kacang-Kacangan Berdasarkan Tekstur Menggunakan Jaringan Syaraf Tiruan', *Jurnal Komputer Terapan*, 6(1), hal. 89-98.
- Godara, N. dan Kumar, S. (2019) 'Opinion Mining using Machine Learning Techniques', *International Journal of Engineering and Advanced Technology (IJEAT)*, 9(2), hal. 4287-4292.
- Gornale, S.S., Basavanna, M. dan Kruthi, R. (2017) 'Fingerprint Based Gender Classification Using Local Binary Pattern', *International Journal of Computational Intelligence Research*, 13(2), hal. 261-271.
- Gungadin, S. (2007) 'Sex determination from fingerprint ridge density', *Internet Journal of Medical Update*, 2(2), hal. 4-7.
- Handayani, F. dan Pribadi, F.S. (2015) 'Implementasi algoritma naive bayes classifier dalam pengklasifikasian teks otomatis pengaduan dan pelaporan masyarakat melalui layanan call center 110', *Jurnal Teknik Elektro*, 7(1), hal. 19-24.
- Kusumawati, R., D'arofah, A. dan Pramana, P.A. (2019) 'Comparison Performance of Naive Bayes Classifier and Support Vector Machine Algorithm for Twitter's Classification of Tokopedia Services', in *Journal of Physics: Conference Series*. IOP Publishing, p. 012016.1-012016.10.
- Siregar, N.C., Siregar, R.R.A. dan Sudirman, M.Y.D. (2020) 'Implementasi Metode Naive Bayes Classifier (NBC) Pada Komentar Warga Sekolah Mengenai Pelaksanaan Pembelajaran Jarak Jauh (PJJ)', *Jurnal Teknologi*, 3(1), hal. 102-110.
- Syahputra, R., Yanris, G. J. dan Irmayani, D. (2022) 'SVM and Naive Bayes Algorithm Comparison for User Sentiment Analysis on Twitter', *Sinkron: Jurnal dan Penelitian Teknik Informatika*, 7(2), hal. 671-678.
- Syaputri, A.W., Irwandi, E. dan Mustakim (2020) 'Naive Bayes Algorithm for Classification of Student Major's Specialization', *Journal of Intelligent Computing and Health Informatics*, 1(1), hal. 15-19.
- Windarti, M.A. dan Suradi, A. (2019) 'Perbandingan Kinerja 6 Algoritme Klasifikasi Data Mining untuk Prediksi Masa Studi Mahasiswa', *Jurnal Telematika Vol*, 1(1), hal. 14-30.
- Zain, F.F. dan Sibaroni, Y. (2019) 'Effectiveness of SVM Method by Naive Bayes Weighting in Movie Review Classification', *Khazanah Informatika: Jurnal Ilmu Komputer dan Informatika*, 5(2), hal. 108-114.

