

BAB II

PEMBAHASAN

2.1 Tinjauan pustaka

Dalam penelitian ini akan digunakan lima tinjauan pustaka yang nantinya dapat mendukung penelitian, berikut merupakan tinjauan pustaka yang diambil yaitu pada tabel dibawah ini.

Tabel 2.1 *Tabel Literatur*

1	Penulis	(Waliyansyah and Fitriyah, 2019)
	Judul	Perbandingan Akurasi Klasifikasi Citra Kayu Jati Menggunakan Metode Naïve Bayes dan K-Nearest Neighbor(K-NN)
	Identifikasi masalah	Menerapkan teknologi untuk membantu dalam menganalisis suatu tekstur kayu jati agar bisa diklasifikasikan ke dalam kelompok-kelompok tertentu
	Metode/Tools	<i>Naïve Bayes Classifier</i> dan <i>K-Nearest Neighbor</i>
	Hasil penelitian	Dari hasil penelitian uji coba klasifikasi citra. Klasifikasi citra kayu jati yang berasal dari sulawesi dengan Metode Naive Bayes paling baik dengan persentase tingkat akurasi sebesar 82,7%. Sedangkan menggunakan metode KNN sebesar 75%. Hal ini dikarenakan akurasi Metode k-NN dipengaruhi oleh banyaknya data latih yang digunakan. Sedangkan Metode Naive bayes dapat menghasilkan akurasi yang baik walaupun memiliki data latih yang sedikit.

Tabel 2.1 Tabel Literatur Lanjutan

2	Penulis	(Putri <i>et al.</i> , 2021)
	Judul	Penerima Manfaat Bantuan Non Tunai Kartu Keluarga Sejahtera Menggunakan Metode NAÏVE BAYES dan KNN
	Identifikasi masalah	Karena masih banyak keluarga tidak mampu lainnya yang belum berkesempatan menerima program bantuan ini. Mengingat banyaknya kriteria dan jumlah masyarakat yang akan menerima KKS, maka akan membuat tim seleksi mengalami kesulitan dalam melakukan penyeleksian.
	Metode/Tools	<i>K-NN</i> , dan <i>Naive Bayes</i>
	Hasil penelitian	hasil Klasifikasi Penerima Manfaat Bantuan Non Tunai Kartu Keluarga Sejahtera Menggunakan Metode Naive Bayes Dan K-Nearst Neighbor dengan performance dari metode algoritma Naive Bayes menghasilkan nilai accuracy sebesar 99.89% dan performance metode algoritma K-Nearst Neighbor 66.46%.

Tabel 2.1 Tabel Literatur Lanjutan

3	Penulis	(Saraswita, 2019)
	Judul	Akurasi Klasifikasi Citra Digital <i>Scenes</i> RGB Menggunakan Model <i>K-Nearest Neighbor</i> dan <i>Naive Bayes</i>
	Identifikasi masalah	pengenalan citra digital pemandangan (<i>scenes</i>) RGB dan akan membandingkan tingkat presentase keberhasilan yang lebih baik diantara metode <i>Naive Bayes</i> dan metode <i>K-Nearest Neighbor (KNN)</i> .
	Metode/Tools	metode <i>Naive Bayes</i> dan metode <i>K-Nearest Neighbor (KNN)</i>
	Hasil penelitian	<i>K-Nearest Neighbor</i> dapat digunakan untuk mengklasifikasi citra digital pemandangan (<i>scenes</i>) RGB lebih baik dari metode <i>Naive Bayes</i> . Akan tetapi tingkat akurasi masing kurang memuaskan, yaitu 76% dan 66%. Maka, untuk penelitian berikutnya metode dapat dikembangkan lagi. Dengan begitu dapat memaksimalkan proses prapengolahan dan ekstraksi ciri citra.

Tabel 2.1 Tabel Literatur Lanjutan

4	Penulis	(Islam <i>et al.</i> , 2019)
	Judul	Klasifikasi Dokumen Tugas Akhir Berbasis Text Mining menggunakan Metode <i>Naïve Bayes Classifier</i> dan <i>K-Nearest Neighbor</i>
	Identifikasi masalah	Pada Program Studi Sistem Informasi UIN SUSKA Riau, penentuan dosen pembimbing tugas akhir ditentukan oleh Ketua Program Studi dengan melakukan kajian terstruktur berdasarkan judul/topik tugas akhir yang diajukan mahasiswa dengan keahlian bidang dosen yang sesuai. Untuk itu, perlu adanya sistem yang dapat memberikan rekomendasi agar pembagian dosen pembimbing lebih efisien.
	Metode/Tools	<i>NBC</i> dan <i>KNN</i>
	Hasil penelitian	Dari percobaan 3 kelas dan 16 kelas diperoleh akurasi terbaik pada percobaan 3 kelas dengan nilai 86,11% untuk <i>Naive Bayes Classifier (NBC)</i> dan 91,67% untuk <i>K-Nearest Neighbor (KNN)</i> dan percobaan 16 kelas dengan nilai 25,20% untuk <i>Naive Bayes Classifier (NBC)</i> dan 22,05% untuk <i>K-Nearest Neighbor (KNN)</i> .

Tabel 2.1 Tabel Literatur Lanjutan

5	Penulis	(Algoritme <i>et al.</i> , 2022)
	Judul	Perbandingan Algoritme Naïve Bayes Dan Knn Terhadap Data Penerimaan Beasiswa (Studi Kasus Lembaga Beasiswa Baznas Jabar)
	Identifikasi masalah	Jumlah pendaftar masuk berlebihan dan mengakibatkan proses seleksi menjadi tidak efisien dan memakan banyak waktu serta pengambilan keputusan yang subjektif
	Metode/Tools	<i>Naïve Bayes</i> Dan <i>K-Nearest Neighbor</i>
	Hasil penelitian	Hasil yang diperoleh adalah algoritma naïve bayes mempunyai akurasi yang sedikit lebih besar dari knn. Hasil uji yang telah dibuat menggunakan algoritma naïve bayes adalah nilai akurasi yang didapat sebesar 80%. Untuk hasil uji yang telah dibuat menggunakan algoritma knn dengan nilai k=7 didapat hasil akurasi yang sebesar 78,97%.

2.1.1 Tinjauan literature 1

Pada literatur yang berjudul Klasifikasi Kayu Jati Menggunakan Metode *Naïve Bayes Dan K-Nearest Neighbor*. Penulis memiliki tujuan yaitu ingin menerapkan teknologi untuk membantu dalam menganalisis suatu tekstur kayu jati agar bisa diklasifikasikan ke dalam kelompok-kelompok tertentu. Proses klasifikasi jati menggunakan pengolahan citra

digital dengan Metode Naive Bayes dan k-Nearest Neighbor (k-NN). Analisis yang digunakan adalah tekstur dengan Metode Gray Level Co-occurrence Matrix (GLCM) dan jarak spasial adalah 1 piksel. Berdasarkan hasil pengujian dan analisis, Metode k- NN secara umum baik dalam mengklasifikasikan 3 jenis kayu jati yaitu semarangan, blora dan sulawesi dengan tingkat akurasi di atas 75%. Akan tetapi klasifikasi paling baik untuk jenis kayu jati sulawesi dengan Metode Naive Bayes, tingkat akurasinya sebesar 82,7%.

2.1.2 Tinjauan literature 2

Pada literatur yang berjudul Penerima Manfaat Bantuan Non Tunai Kartu Keluarga Sejahtera Menggunakan Metode NAÏVE BAYES dan KNN. Penulis memiliki tujuan yaitu untuk mengklasifikasikan penerima manfaat Kartu Keluarga Sejahtera, karena masih banyak keluarga yang tidak mampu lainnya yang belum berkesempatan menerima program bantuan. Berdasarkan tujuan penelitian ini untuk dapat mengetahui performance metode algoritma Naïve Bayes dalam mengklasifikasi data penerima manfaat bantuan non tunai kartu keluarga sejahtera di kecamatan Harjamukti. hasil Klasifikasi Penerima Manfaat Bantuan Non Tunai Kartu Keluarga Sejahtera Menggunakan Metode Naïve Bayes Dan K-Nearst Neighbor dengan performance dari metode algoritma Naïve Bayes menghasilkan nilai accuracy sebesar 99.89% dan performance metode algoritma K-Nearst Neighbor 66.46%.

2.1.3 Tinjauan literature 3

Pada literatur yang berjudul Akurasi Klasifikasi Citra Digital Scenes RGB Menggunakan Model *K-Nearest Neighbor* dan *Naive Bayes*. Penulis memiliki tujuan yaitu penelitian ini ingin membuktikan berapa presentase keberhasilan pengenalan citra digital pemandangan (*scenes*) RGB dan ingin membandingkan tingkat nilai presentase keberhasilan yang lebih baik antara metode *Naive Bayes* dan metode *K-Nearest Neighbor (KNN)*. Peneliti mengambil data set gambar yang digunakan yaitu berjumlah 8000 gambar pemandangan gambar sebanyak 948 gambar dalam 3 kategori, yaitu pesisir sebanyak 360 gambar, hutan sebanyak 328 gambar, jalan raya sebanyak 260 gambar. Kemudian menentukan jenis-jenis gambarnya. Setelah itu melakukan klasifikasi yaitu dengan membagi koleksi data menjadi dua bagian utama yaitu data latih dan data uji terlebih dahulu. Pada pengklasifikasian ini peneliti menggunakan fungsi jarak dari data baru ke data training (kedekatan). Prinsip kerja *K-Nearest Neighbor (KNN)* adalah dengan mengambil atau mencari jarak terdekat antara data yang akan dievaluasi dengan K tetangga (*neighbor*) terdekatnya dalam data pelatihan. Kemudian menggunakan metode *Naive Bayes* dan memperkirakan semua atribut independen atau tidak saling berketergantungan yang diberikan oleh nilai pada variabel. Untuk menguji tingkat keberhasilannya, peneliti menggunakan model perhitungan. Hasil menyatakan bahwa nilai akurasi terhadap klasifikasi metode *K-Nearest Neighbor (KNN)* sebesar 76% dan metode *Naive Bayes* sebesar 60%.

2.1.4 Tinjauan literature 4

Pada literatur yang berjudul Klasifikasi Dokumen Tugas Akhir Berbasis Text Mining menggunakan Metode *Naïve Bayes Classifier* dan *K-Nearest Neighbor*. Penulis memiliki tujuan yaitu untuk mempermudah Kaprodi dalam menentukan dosen pembimbing bagi mahasiswa yang mengajukan Proposal Tugas Akhir. Dataset yang digunakan pada penelitian ini yaitu data Tugas Akhir (TA) program studi Sistem Informasi Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim (UIN SUSKA) Riau tahun 2013 sampai tahun 2018. Total data keseluruhan berjumlah sebanyak 480 judul TA. Menentukan atribut, kemudian setelah dataset diperoleh dan atribut ditentukan selanjutnya melakukan pembersihan pada data atau *cleaning* data. Kemudian eliminasi data yang tidak sesuai menggunakan text *Preprocessing*. Klasifikasi NBC menggunakan modul scikit-learn pada python yaitu Multi Nomial NB dan klasifikasi KNN menggunakan modul scikit-learn pada python yaitu *K Neighbors Classifier*. Pada perhitungan *KNN*, peneliti ini melakukan percobaan nilai k menggunakan nilai k=6, 7, 8, 9 dan 10 dan didapatkan hasil k=7 terbaik. Pada algoritma *KNN* nilai k yang memiliki nilai presentase akurasi tertinggi yaitu k=6, k=7, k=8 dan k=9, dengan nilai k=7 sebesar 91,67% dan sebesar 86,11% untuk nilai akurasi yang sama pada nilai k lainnya. Akurasi yang bagus dapat diperoleh jika sebaran data disetiap kelas tidak memiliki range yang terlalu jauh seperti pada percobaan menggunakan 3 kelas.

2.1.5 Tinjauan literature 5

Pada literatur yang berjudul Perbandingan Algoritme Naïve Bayes Dan Knn Terhadap Data Penerimaan Beasiswa (Studi Kasus Lembaga Beasiswa Baznas Jabar). Penulis memiliki tujuan dengan menggunakan algoritma klasifikasi data mining terhadap data yang telah ada akan diketahui pola pemetaan karakteristik pemohon ke keputusan yang diambil sehingga proses seleksi akan menjadi lebih mudah, konsisten dan dapat menghindari unsur subjektifitas dari pengambil keputusan.. Hasil yang diperoleh adalah algoritma naïve bayes mempunyai akurasi yang sedikit lebih besar dari knn. Hasil uji yang telah dibuat menggunakan algoritma naïve bayes adalah nilai akurasi yang didapat sebesar 80%. Untuk hasil uji yang telah dibuat menggunakan algoritma knn dengan nilai $k=7$ didapat hasil akurasi yang sebesar 78,97%.

2.2 Tapis lampung

Kain tenun tapis merupakan perangkat yang memiliki makna beraneka ragam yang berhubungan dengan kepercayaan, perasaan sakral dan pemuasan akan cita rasa keindahan (Isbandiyah and Supriyanto, 2019). Kain tapis ini kain ciri khas di daerah Lampung yang memiliki bentuk yang menyerupai sarung dan biasa digunakan oleh wanita suku Lampung. Kain ini dibuat dengan cara di tenun dengan menggunakan benang kapas dengan motif bahan sugi, benang perak atau benang emas dengan cara disulam. Pada zaman dahulu para gadis di Lampung diharuskan bisa menyulam sejak dini, kemudian pada saat ini para gadis sangat jarang sekali ditemui yang mahir dalam menyulam. Kain tapis biasa

hanya digunakan pada acara – acara tertentu yang menjadikannya kain yang sakral karena digunakan pada acara-acara yang sakral. Bentuk-bentuk khas pada tapis yaitu seperti alam, flora, fauna dan kehidupan masyarakat Lampung. Kain tapis memiliki banyak macam, pada penelitian ini berfokus pada jenis motif lereng dan jung sarat.

2.2.1 Tapis Motif Sasab

Motif sasab melambangkan kerapatan yang artinya kerapatan antara keluarga laki-laki dan keluarga wanita. Motif sasab berarti penuh atau padat mengandung makna ilmu yang bermanfaat lahir batin, sesuai ketentuan agama atau kepercayaan yang dianut.



Gambar 2. 1 *Tapis Motif Sasab*

2.2.2 Tapis Sasab Ketik

Tapis sasab ketik yaitu perpaduan antara motif sasab dan motif ketik. Memiliki makna simbolis sebagai lambang kesucian yang dapat melindungi pemakainya dari segala kotoran dari luar.



Gambar 2. 2 *Tapis Motif Sasab Ketik*

2.2.3 Tapis Motif Wajik

Merupakan sarung pakan lungsi. Berbahan dasar benang kapas berlajur horizontal, dan warna merah hati, hitam,serta kuning. Ragam hias disulam memakai benang emas. . Motifnya, yaitu seperti belah ketupat atau wajik.



Gambar 2. 3 *Tapis Motif Wajik*

2.3 Citra

Citra dapat diartikan sebagai hasil penangkapan suatu objek yang mempunyai nilai koordinat spesial, dan memiliki intensitas kecerahan warna pada suatu objek (Zahra *et al.*, 2021). Citra merupakan bentuk informasi visual sehingga banyak informasi yang dapat digali dari sebuah citra (Rianto P and Harjoko A, 2017). Citra yang digunakan dalam penelitian ini adalah kain tapis Lampung. Pengambilan citra ini digunakan sebagai data dalam penelitian ini. Data ini kemudian akan diproses menggunakan metode *KNN* dan *Naives Bayes*. Citra digital merupakan citra yang dapat diolah atau diproses oleh komputer. Citra tidak selalu sempurna, citra juga perlu diperbaiki atau disempurnakan. Untuk memperbaiki citra yang rusak seperti citra warna yang kurang tajam, citra yang kabur, kurang terang dan lainnya maka diperlukan pengolahan citra digital untuk mengolahnya. Karena jika citra kurang sempurna maka

informasi yang akan didapat pada citra akan kurang atau tidak relevan. Dalam pengolahan citra objek yang digunakan adalah sebuah citra dari sebuah objek tertentu yang mana citra tersebut mengandung sejumlah elemen dasar. Elemen dasar tersebut di manipulasi dalam pengolahan citra, elemen tersebut adalah:

1. Warna yang diterima oleh mata merupakan hasil kombinasi cahaya dengan panjang gelombang berbeda. Kombinasi warna yang memberikan rentang warna yang paling lebar adalah *red* (R), *green* (G), *blue* (B).
2. Kecerahan juga disebut sebagai intensitas cahaya. Kecerahan terhadap sebuah piksel (titik) di dalam citra bukanlah intensitas yang nyata, tetapi sebenarnya adalah intensitas rata-rata dari suatu area yang melingkupinya.
3. Kontras (*Contrast*) merupakan sebuah sebaran terang dan gelap di dalam sebuah gambar. Citra dengan kontras rendah dicirikan oleh sebagian besar komposisi citranya adalah sebagian besar gelap.
4. Tekstur diartikan sebagai derajat keabuan di dalam sekumpulan piksel-piksel yang bertetangga. Jadi tekstur tidak dapat didefinisikan untuk sebuah piksel. Sistem visual manusia menerima informasi citra sebagai suatu kesatuan. Resolusi citra yang diamati ditentukan oleh skala dimana tekstur tersebut dipersepsi.

2.4 Pengolahan Citra Digital

Pengolahan Citra Digital pada dasarnya berfungsi untuk memperbaiki kualitas gambar dilihat dari peningkatan kontras, transformasi warna, restorasi dan citra. Pengolahan citra yaitu cara mengolah citra dengan berbagai teknik tertentu, seperti *preprocessing*, *resize*, *grayscale*, *thresholding*, ekstraksi dan normalisasi. Pengolahan citra dapat digunakan untuk memperbaiki citra yang rusak. Pengolahan citra dan pengenalan pola menjadi bagian dari proses pengenalan citra untuk mendapatkan ciri khas dari suatu citra yang hendak dikenali. Jadi secara umum pengolahan citra digital merupakan langkah-langkah teknik dalam mengestimasi ciri-ciri objek di dalam citra, pengukuran ciri yang berkaitan dengan geometri objek dan menginterpretasi geometri tersebut (Rianto P and Harjoko A, 2017).

2.4.1 Preprocessing

Preprocessing adalah sebuah tahapan dimana citra hasil acquisition di perbaiki dan disesuaikan agar hasil pattern analysis dan decision yang dilakukan semakin cepat dan akurat (Noercholis, Muslim and Maftuch, 2013). Perbedaan data ciri yang diharapkan menuntut adanya perbedaan tahapan pada *preprocessing* pengujian, citra masukan seperti halnya pada pengujian kain tapis Lampung ini. Akan melalui tahapan *resize*, *grayscale*, *thresholding*, dan ekstraksi selanjutnya yaitu normalisasi. Tahap *preprocessing* dilakukan untuk memperbaiki dan menyesuaikan citra.

2.4.2 Resize

Dalam proses pengubahan ukuran (*resizing*), metode yang digunakan adalah interpolasi tetangga terdekat. Pada interpolasi ini, nilai keabuan suatu titik piksel diambil dari nilai keabuan pada titik asal yang paling dekat dengan koordinat hasil perhitungan dari transformasi spasial.

2.4.3 Grayscale

Suatu citra digital grayscale adalah suatu citra dimana nilai dari setiap piksel merupakan sampel tunggal. Citra yang ditampilkan dari citra jenis ini terdiri atas warna abu-abu, bervariasi pada warna hitam pada bagian yang intensitas terlemah dan warna putih pada intensitas terkuat. Tujuan konversi ini agar lebih memudahkan dalam melakukan analisis citra yang diperlukan pada proses tahapan pengolahan citra selanjutnya. Proses grayscale mengurangi dimensi yang dimiliki oleh citra, dengan melakukan perubahan citra tiga warna menjadi hanya satu warna yaitu warna keabuan.

2.4.4 Thresholding

Thresholding merupakan suatu metode untuk mengubah citra skala keabuan menjadi citra biner. Dalam metode ini diharuskan memiliki ketepatan dalam nilai ambang batas. Melakukan pengaturan atau manipulasi nilai ambang agar citra dapat tersegmentasi. Setelah melakukan proses *grayscale*, maka dapat menggunakan teknik ini yaitu dengan merubah warna citra menjadi hitam putih.

2.5 Ekstraksi Citra

Proses ekstraksi dilakukan untuk mendapatkan ciri tertentu dari suatu citra digital. Ciri itu berupa informasi dari citra objek yang dapat membuat agar dapat dikenali, dikelompokkan atau dibedakan dari citra lainnya. Kemudian setelah mendapatkan informasi pada citra, maka informasi tersebut akan dijadikan sebagai parameter atau nilai untuk membedakan antara citra satu dengan yang lainnya. Untuk dapat membedakan antara objek yang satu dengan yang lainnya dapat menggunakan proses klasifikasi atau identifikasi. Ekstraksi ciri merupakan proses pengambilan ciri atau karakteristik objek yang dapat digunakan sebagai pembeda dari objek-objek lainnya. Tiap-tiap citra atau tiap-tiap objek memiliki nilai-nilai yang merupakan ciri unik jika tiap-tiap citra di ekstraksi. *GLCM* yaitu di mana matriks yang merepresentasikan banyaknya suatu pixel i dan pixel tetangga j yang berada pada sebuah citra. Ekstraksi fitur dari *GLCM* (*Gray-Level Cooccurrence Matrix*) menghasilkan fitur yaitu *Contrast*, *Energy* dan *Homogeneity*, *Correlation*. *GLCM* memiliki 4 arah sudut dalam ketetanggaan yaitu 0° , 45° , 90° , 135° . Sudut 0° ketetanggaan dihitung dengan jarak satu pixel ke kanan, sudut 45° ketetanggaan dihitung dengan jarak satu piksel ke kanan atas. Sudut 90° ketetanggaan dihitung dengan jarak satu piksel ke atas. Sudut 135° ketetanggaan dihitung jarak satu piksel ke kiri atas.

a) Contrast

Contrast merupakan fitur yang menampilkan perbedaan tingkat warna atau skala keabuan (*grayscale*) yang muncul pada sebuah citra. Jika piksel ketetanggaannya mempunyai nilai yang sama *Contrast* akan bernilai 0.

b) Energy

Energy merupakan ukuran keseragaman pada citra. Semakin tinggi kemiripan semakin tinggi pula nilai *Energy*.

c) Homogeneity

Homogeneity merupakan ukuran kebersamaan. *Homogeneity* akan bernilai tinggi jika semua piksel mempunyai nilai yang uniform.

d) Correlation

Merupakan keterkaitan linear dari derajat dari citra keabuan. *Correlation* bernilai antara -1 hingga 1.

2.6 Normalisasi

Tahap normalisasi merupakan tahap persiapan untuk tahap identifikasi, yaitu membuat data menjadi ukuran tetap. Yang nantinya data tersebut diproses untuk mendapatkan cirinya melalui ekstraksi ciri. Dengan menggunakan *min-max normalization* melakukan transformasi linier terhadap data asli sehingga menghasilkan keseimbangan nilai perbandingan antar data saat sebelum dan sesudah proses.

2.7 Klasifikasi

Klasifikasi yaitu suatu proses pengkategorian atau pengelompokan yang dilakukan terhadap sekumpulan data. Klasifikasi dilakukan dengan membagi koleksi data menjadi dua bagian utama yaitu data latih dan data uji

terlebih dahulu. Dari data latih di hitung dengan menggunakan metode tertentu kemudian, diperoleh model klasifikasi yang nantinya akan digunakan untuk penentuan kelas terhadap data uji (Saraswita, 2019). Klasifikasi merupakan metode untuk memprediksi nilai dari sekelompok atribut.

2.8 Metode KNN

K-Nearest Neighbor (KNN) adalah algoritma sederhana yang menyimpan semua case yang tersedia dan mengklasifikasikan data atau case baru berdasarkan ukuran kesamaan. Itu sebagian besar digunakan untuk mengklasifikasikan titik data berdasarkan bagaimana tetangganya diklasifikasikan (Pola *et al.*, no date). *KNN* cenderung berfungsi paling baik pada kumpulan data yang lebih kecil yang tidak memiliki banyak fitur. Penentuan jumlah kelompok (k) dapat dilakukan dengan berbagai metode, antara lain ditentukan secara acak, atau ditentukan secara pasti (Iswari, Wella and Ranny, 2017). Klasifikasi menggunakan metode *KNN* ini cara menggunakannya yaitu dengan jarak yang paling dekat dengan objek. Hal ini hanya didasarkan pada gagasan bahwa suatu objek yang dekat satu sama lain juga akan memiliki karakteristik yang mirip.

Algoritma KNN :

1. Mulai
2. Tentukan parameter K
3. Hitung jarak antara data yang akan dievaluasi dengan semua pelatihan Jarak ini didapatkan dari data yang telah di ekstraksi metode Gray Level Co-Ocurece Matrix yaitu data baru yang akan di uji di kurangi dengan semua

data pelatihan yang di sebut dengan jarak Euclidian. Persamaan euclidean distance :

$$Distance = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (2.1)$$

Dimana,

- a) $d(x,y)$ adalah jarak antara data x ke data y
 - b) x_i adalah data testing ke-i
 - c) y_i adalah data training ke-i
 - d) n adalah dimensi data
4. Urutkan jarak yang terbentuk (urut naik) Setelah diketahui jarak setiap data latih, maka data di urutkan atau di sorting mulai dari data terkecil sampai data yang paling besar lalu cari jarak yang paling kecil (terdekat) dengan data latih.
 5. Pasangkan kelas yang bersesuaian Setelah di tentukan K terdekatnya berapa, maka bisa di ketahui beberapa kelas yang akan di tujukan pada data uji tersebut.
 6. Cari jumlah kelas dan tetangga yang terdekat dan tetapkan kelas tersebut sebagai kelas data yang akan dievaluasi menggunakan kategori Nearest Neighbor yang paling mayoritas.
 7. Selesai.

KNN memiliki beberapa kelebihan yaitu ketangguhan terhadap data latih yang memiliki *noise* dan efektif apabila *training* datanya besar.

2.9 Naive Bayes

Algoritma *Naive Bayes* adalah salah satu algoritma yang terdapat pada teknik klasifikasi. *Naive Bayes* dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya sehingga disebut sebagai Teorema Bayes (Alita, Sari and Rahman Isnain, 2021). Prinsip umumnya adalah mengasumsikan bahwa nilai suatu atribut tidak bergantung dan mempengaruhi atribut yang lainnya (Waliyansyah and Fitriyah, 2019). Klasifikasi *Naive Bayes* adalah model yang paling sederhana, karena mengasumsikan bahwa semua atribut adalah independen (Iswari, Wella and Ranny, 2017). Ciri utama dari *Naive Bayes* ini adalah asumsi yang sangat kuat akan independen dari masing-masing kondisi / kejadian. *Naive bayes* merupakan metode pengklasifikasian berdasarkan probabilitas sederhana dan dirancang agar dapat dipergunakan dengan asumsi antar variabel penjelas saling bebas (*independen*). Metode ini sering digunakan dalam menyelesaikan masalah dalam bidang pembelajaran karena metode ini dikenal memiliki tingkat akurasi yang tinggi dengan perhitungan sederhana. Persamaan *Naive Bayes* sebagai berikut :

$$P(H_i|X) = \frac{P(X|H_i)P(H_i)}{P(X)} \quad (2.2)$$

Dimana :

H_i = Hipotesis data X yang merupakan kelas yang spesifik

X = Data dengan kelas yang belum diketahui

$P(H_i|X)$ = Probabilitas hipotesis H_i berdasarkan kondisi X

$P(X|H_i)$ = Probabilitas x berdasarkan kondisi pada hipotesis H_i .

$P(H_i)$ = Probabilitas hipotesis H_i .

$P(X)$ = Probabilitas dari data X .

Metode yang diusulkan mempertimbangkan input data set dengan nilai atribut numerik dan distribusi *Gaussian*. Untuk distribusi *Gaussian mean* (μ) pada persamaan (2.3) dan standar *deviasi* (σ) perlu dihitung yang dengan persamaan (2.4) dibawah ini:

$$\mu = \frac{X_1+x_2+\dots+X_n}{n} \quad (2.3)$$

$$\sigma = \frac{\sum_{i=1}^n (X_i - \mu)^2}{n} \quad (2.4)$$

Sedangkan fungsi distribusi *Gaussian* dinyatakan dalam persamaan (2.5)

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2(\sigma^2)}} \quad (2.5)$$

Dimana,

μ = Hasil perhitungan μ tiap atribut data latih berdasarkan kelas yang sudah dilakukan sebelumnya.

σ = Hasil perhitungan σ tiap atribut data latih berdasarkan kelas yang telah dilakukan sebelumnya.

X = Nilai atribut data uji yang digunakan.

π = 3,14

e = 2,718282