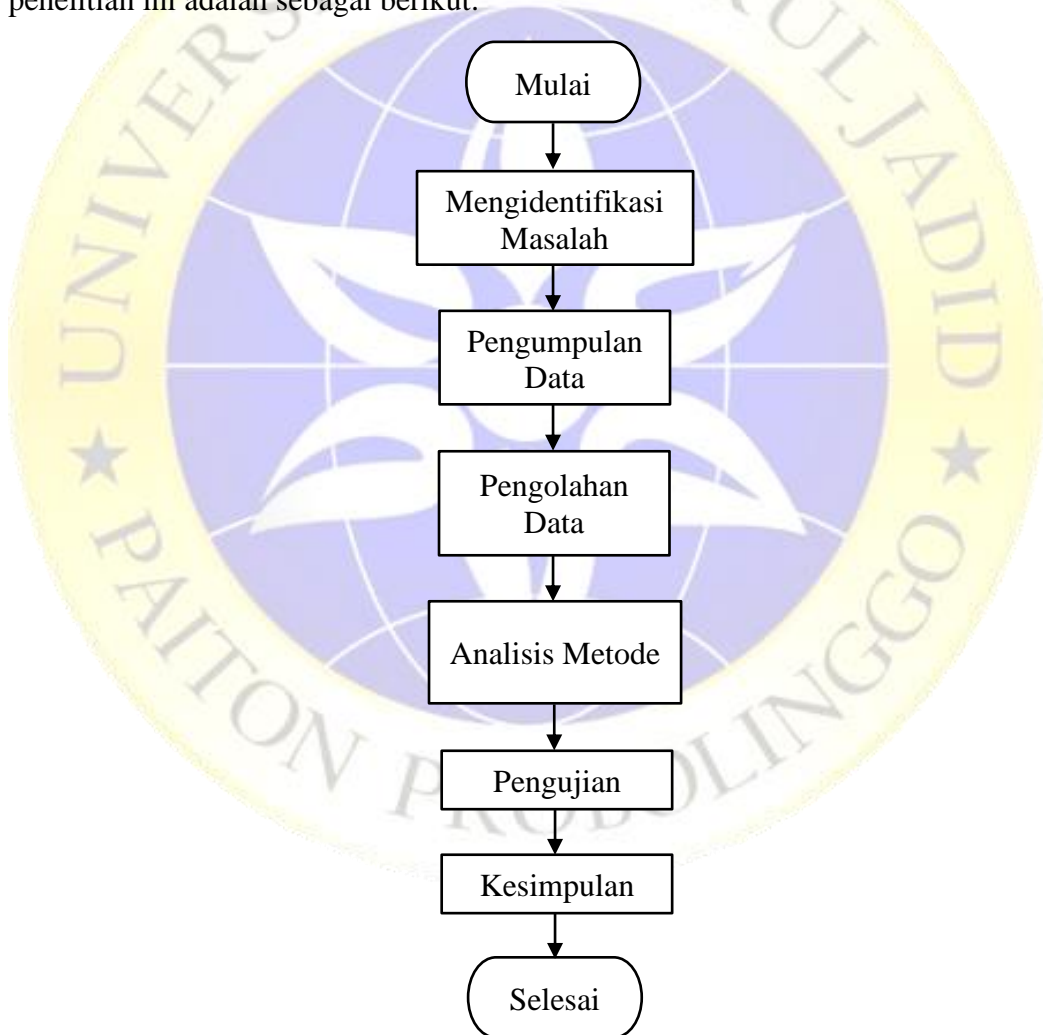


BAB III METODE PENELITIAN

3.1 Kerangka Penelitian

Dalam melakukan proses penelitian, perlu adanya sebuah kerangka penelitian. Di mana kerangka penelitian merupakan suatu konsep atau rancangan yang disusun secara berurutan dan saling berhubungan sehingga penelitian yang dilakukan dapat terselesaikan dengan baik. Adapun kerangka penelitian dari penelitian ini adalah sebagai berikut:



Gambar 3. 1 Kerangka Penelitian

Dari Gambar 3.1 di atas dapat dilihat bagaimana alur atau tahapan dari penelitian ini. Di mana pada tahap pertama ialah mengidentifikasi masalah, identifikasi masalah yang dimaksud adalah mendefinisikan masalah yang

ditemukan sebagai langkah awal adanya penelitian ini. Setelah proses identifikasi masalah ditemukan tahap selanjutnya ialah proses pengumpulan data atau kebutuhan yang dibutuhkan dalam penelitian ini. Setelah data yang dibutuhkan diperoleh, data tersebut diolah sehingga menjadi data yang siap untuk diimplementasikan terhadap metode yang digunakan. Analisis metode ialah proses mengimplementasikan metode yang digunakan, di mana dalam penelitian ini metode yang digunakan adalah metode PCA untuk mereduksi atribut dari dataset, dan metode KNN untuk mengklasifikasikan data jamur. Tahap selanjutnya ialah proses pengujian, tahap ini dilakukan untuk mengetahui hasil dari analisis metode tersebut. Setelah hasil pengujian diperoleh maka dapat ditarik kesimpulan mengenai seberapa baik kinerja metode yang digunakan. Berikut penjabaran lengkap mengenai alur atau tahapan dari kerangka penelitian ini.

3.2 Model Pengembangan

3.2.1 Pengumpulan Data

Data yang digunakan dalam penelitian ini merupakan data sekunder, di mana data yang digunakan dalam penelitian ini tidak diperoleh secara langsung dari sumbernya, melainkan diperoleh dari laman internet <https://archive.ics.uci.edu/ml/datasets/mushroom> (UCI Machine Learning Repository). Data yang disumbangkan oleh Jeff Schlimmer pada tahun 1987 ini berjumlah sebanyak 8124 data dengan 22 atribut meliputi bentuk topi, permukaan topi, warna topi, membran, bau, lampiran lamela, jarak lamela, ukuran lamela, warna lamela, bentuk tangkai, akar tangkai, permukaan tangkai di atas, permukaan tangkai di bawah, warna tangkai di atas cincin, warna tangkai di bawah cincin, tipe kerudung, warna kerudung, jumlah cincin, tipe cincin, warna spora, populasi, dan habitat. Data sebanyak 8124 tersebut dikelompokkan ke dalam dua kelas yakni kelas jamur pangan dan kelas jamur beracun, sebanyak 3916 jamur termasuk ke dalam jamur beracun dan sebanyak 4208 jamur termasuk ke dalam kelas jamur pangan atau layak konsumsi.

3.2.2 Pengolahan Data

Setelah data yang dibutuhkan tersedia, langkah selanjutnya adalah mengolah data tersebut sehingga menjadi data yang siap diterapkan ke dalam metode yang akan diimplementasikan. Proses pengolahan data yang akan dilakukan dalam penelitian ini adalah mengubah nilai dari atribut dataset yang masih berupa inisial menjadi sebuah angka. Angka tersebut dimulai dari angka 1 sampai seterusnya sebanyak jumlah nilai dari atribut tersebut. Jika ada atribut yang memiliki nilai hilang maka diganti dengan angka 0. Berdasarkan analisa yang telah dilakukan pada dataset ini terdapat satu atribut yaitu akar tangkai yang memiliki nilai hilang sebanyak 2480. Kemudian atribut dengan nama tipe kerudung pada penelitian ini tidak digunakan karena nilai dari atribut tersebut tidak bervariasi atau hanya memiliki satu nilai. Sehingga atribut yang awalnya terdapat 22 atribut menjadi 21 atribut karena atribut tipe kerudung tidak digunakan. Pada tahap ini juga dilakukan pembagian data dari data observasi. Di mana data observasi yang berjumlah 8124 data dibagi ke dalam dua data, yaitu data *training* sebanyak 75% dari data observasi dan sisanya sebanyak 25% dari data observasi dijadikan sebagai data *testing*.

3.2.3 Analisis Metode

Metode yang digunakan dalam penelitian ini adalah metode PCA dan KNN. Sesuai dengan fungsinya algoritma PCA akan melakukan reduksi atribut terhadap data yang digunakan. Melihat atribut dalam dataset penelitian ini yang terbilang cukup banyak yakni sebanyak 21 atribut, sehingga perlu disederhanakan dengan melakukan reduksi atribut tanpa harus kehilangan banyak informasi dari atribut awal. Setelah hasil dari reduksi atribut menggunakan metode PCA didapatkan, langkah selanjutnya ialah mengimplementasikan data jamur yang atributnya telah direduksi tersebut ke dalam algoritma klasifikasi KNN. Penerapan algoritma KNN dilakukan untuk mengklasifikasikan jamur ke dalam dua kelas yakni kelas jamur beracun dan kelas jamur pangan atau layak konsumsi.

a. Principal Component Analysis (PCA)

Berdasarkan tahapan algoritma PCA yang telah dijabarkan pada landasan teori di atas, tahap pertama yang harus dilakukan adalah menghitung varians dan kovarian dari atribut dataset menggunakan persamaan 2.1 dan 2.2 sehingga diperoleh matrik kovarian berukuran $n \times n$. Perhitungan nilai varians dan kovarian digunakan untuk mengetahui korelasi antar atribut.

Misal ada 3 atribut yang terdiri dari atribut x , y , dan z , maka matrik kovarian yang akan terbentuk adalah matrik berukuran 3×3 , yaitu:

$$\begin{bmatrix} cov(x,x) & cov(x,y) & cov(x,z) \\ cov(y,x) & cov(y,y) & cov(y,z) \\ cov(z,x) & cov(z,y) & cov(z,z) \end{bmatrix}$$

Di mana $cov(x,x)$, $cov(y,y)$, dan $cov(z,z)$ nilainya sama dengan perhitungan varians dari atribut x , y , dan z . Sehingga matrik kovariannya sebagai berikut:

$$\begin{bmatrix} var(x) & cov(x,y) & cov(x,z) \\ cov(y,x) & var(y) & cov(y,z) \\ cov(z,x) & cov(z,y) & var(z) \end{bmatrix}$$

Setelah matrik kovarian didapat, selanjutnya adalah mencari *eigenvalues* dari matrik kovarian tersebut dengan menggunakan persamaan 2.3. Nilai *eigen* dibentuk dari diagonal matrik kovarian ($cov(x,x)$, $cov(y,y)$, $cov(z,z)$). Kemudian nilai *eigen* yang sudah didapatkan diurutkan dari yang terbesar hingga yang terkecil ($\lambda_1, \lambda_2, \lambda_3$).

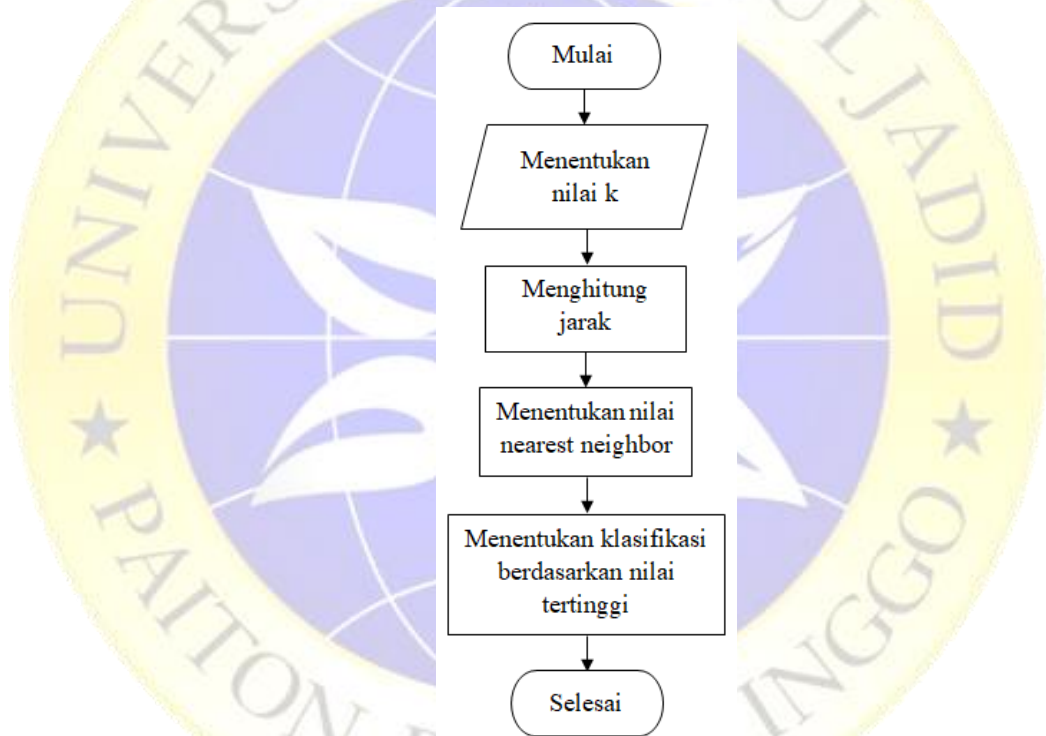
Tahap selanjutnya adalah memilih beberapa *principal component* yang memiliki korelasi paling besar dengan menggunakan persamaan 2.4, varians kovarian didapat dari penjumlahan total diagonal matrik kovarian atau nilai $eigen(\lambda_1 + \lambda_2 + \lambda_3)$. Sehingga didapat beberapa *principal component* dengan proporsi varians sebesar yang diinginkan.

Setelah didapatkan sejumlah *principal component* dengan persentase sebesar yang diinginkan, selanjutnya adalah melakukan perhitungan *eigenvector*. Perhitungan *eigenvector* ini digunakan untuk menentukan atribut apa yang memiliki korelasi paling besar dari *principal component* yang telah dipilih, untuk menghitung bobot atribut berdasarkan *eigenvector* dapat menggunakan persamaan 2.5. penentuan atribut yang memiliki korelasi paling

besar dapat dilihat berdasarkan perbandingan nilai korelasi pada tabel matrik bobot atribut.

b. *K-Nearest Neighbor* (KNN)

Setelah proses pereduksian atribut didapatkan, langkah selanjutnya adalah mengklasifikasikan jenis jamur berdasarkan atribut baru yang diperoleh dari hasil reduksi atribut menggunakan PCA. Metode yang digunakan untuk mengklasifikasikan jenis jamur pangan dan jamur beracun tersebut adalah metode KNN. Langkah-langkah untuk melakukan klasifikasi menggunakan metode KNN dapat dilihat pada gambar di bawah ini:



Gambar 3. 2 Tahapan KNN

Sumber: (Zubair & Muslikh, 2017)

Langkah pertama dalam mengimplementasikan algoritma KNN adalah menentukan nilai K terlebih dahulu. Setelah nilai K ditentukan, tahap selanjutnya ialah menghitung jarak antara data latih dengan data uji dengan menggunakan metode *euclidean* dan *manhattan distance* . Dari hasil perhitungan jarak yang didapatkan maka dapat diketahui nilai tetangga terdekat dari data uji. Tahap terakhir adalah mengklasifikasikan data uji tersebut berdasarkan nilai tertinggi dari tetangga terdekat.

3.2.4 Pengujian

Proses pengujian dalam penelitian ini menggunakan model *confusion matrix*. Di mana tahap ini dilakukan dengan tujuan mengetahui tingkat keberhasilan dari penerapan metode PCA dan metode KNN untuk mengklasifikasikan jamur bercun dan dapat dikonsumsi dari dataset yang digunakan. Tingkat keberhasilan suatu penerapan metode dapat dilihat dari hasil akurasi yang didapatkan. Dalam penelitian ini pengujian dilakukan berdasarkan data *testing* yang telah ditentukan sebanyak 25% dari data observasi. Untuk menghitung nilai akurasi dapat menggunakan persamaan 2.8.

3.2.5 Kesimpulan

Setelah hasil akurasi diperoleh maka dapat ditarik kesimpulan berdasarkan hasil akurasi tersebut. Jika hasil dari perhitungan akurasi yang didapatkan bernilai tinggi maka dapat disimpulkan bahwa penerapan metode PCA dan KNN sangat baik dalam mengklasifikasikan jamur yang dapat dikonsumsi dan jamur beracun. Begitupun sebaliknya, jika nilai akurasi rendah maka kedua metode tersebut dinilai kurang baik dalam mengklasifikasikan jamur pangan dan jamur beracun.