



Hak cipta dan penggunaan kembali:

Lisensi ini mengizinkan setiap orang untuk mengubah, memperbaiki, dan membuat ciptaan turunan bukan untuk kepentingan komersial, selama anda mencantumkan nama penulis dan melisensikan ciptaan turunan dengan syarat yang serupa dengan ciptaan asli.

Copyright and reuse:

This license lets you remix, tweak, and build upon work non-commercially, as long as you credit the origin creator and license it on your new creations under the identical terms.

BAB 3

METODOLOGI PENELITIAN

Metodologi penelitian akan dijabarkan sesuai dengan langkah-langkah yang hendak dilakukan dalam penyusunan dan mengerjakan penelitian. Langkah-langkah penelitian yang dijabarkan dimulai dari awal hingga akhir selesai (dokumentasi). Beberapa contohnya yang dapat diikuti antara lain pengumpulan data, pengambilan sampel, analisis kebutuhan, perancangan, implementasi, pengujian, evaluasi, serta dokumentasi.

3.1 Studi Literatur

Dalam tahapan ini, akan mempelajari teori-teori yang digunakan. Caranya dengan membaca jurnal serta artikel-artikel dari sumber yang terpercaya.

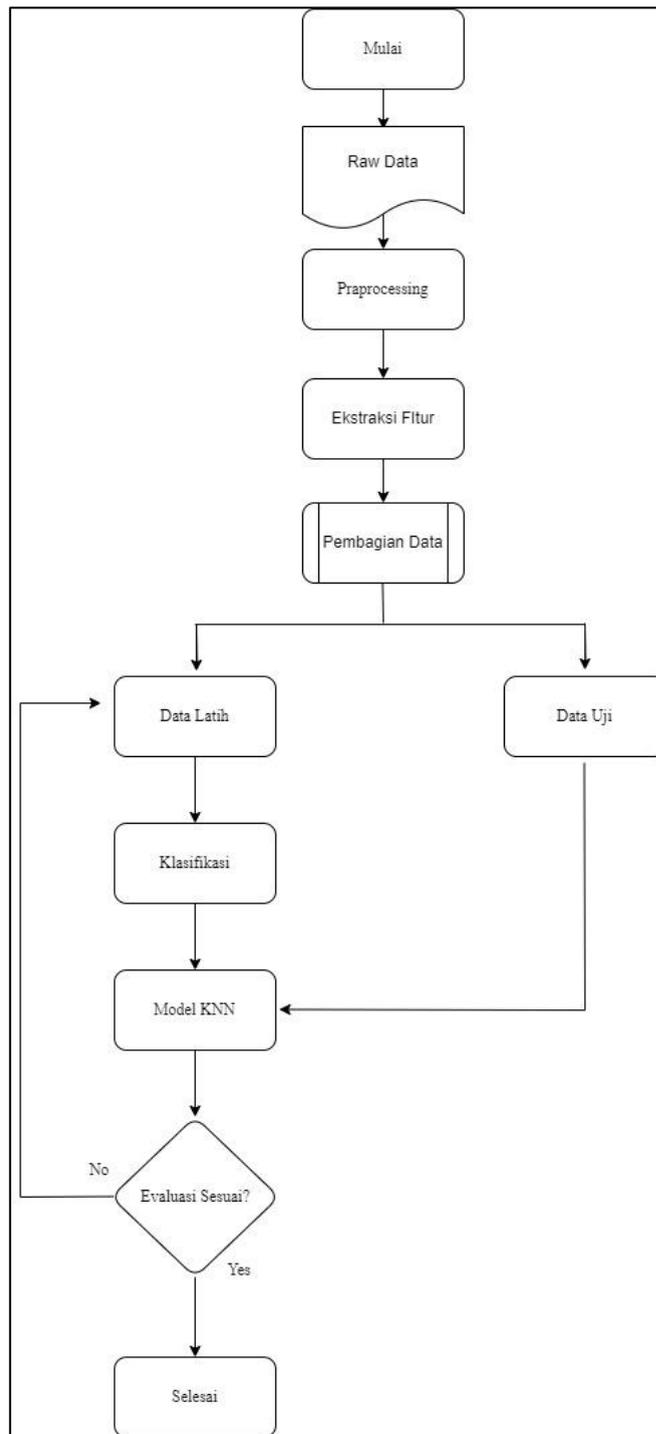
3.2 Pengumpulan Dataset

Pada tahap selanjutnya, dilakukan pengumpulan dataset untuk mendapatkan data klasifikasi bidang pekerjaan.

3.3 Prosedur Penelitian

Selanjutnya, akan dilakukan dalam beberapa tahapan. Berikut penjelasan ada pada Gambar 3.1 dibawah ini.





Gambar 3.1 Flowchart tahapan penelitian

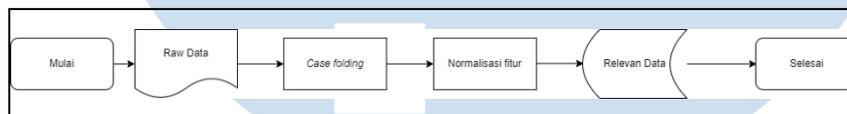
UNIVERSITAS
MULTIMEDIA
NUSANTARA

1. Raw Data

Raw data yang dipakai untuk penelitian yaitu, data yang terdapat di website Kaggle.

2. Praprocessing

Dalam tahap ini, akan dilakukan praprocessing. Praprocessing merupakan proses untuk mempersiapkan data mentah yang belum diolah sebelum dilakukan proses lain. Pada umumnya, praproses data dilakukan dengan cara mengeliminasi data yang tidak sesuai atau mengubah data menjadi bentuk yang lebih mudah diproses oleh sistem [13].



Gambar 3.2 Flowchart tahapan praprocessing

3. Ekstraksi fitur

Ekstraksi fitur adalah proses pengambilan ciri pada sebuah objek yang dapat menggambarkan karakteristik dari masing – masing objek.

4. Pembagian Data

Data yang dibagikan merupakan data proses dari pengambilan ciri sebuah objek yang dapat menggambarkan karakteristik dari masing – masing objek. Pada tahap ini, akan dilakukan pembagian data sesuai dengan K-Nearest Neighbors yang diinginkan ke dalam data *training* (data latih) beserta data uji.

a. Data Latih

Data latih adalah yang akan dipakai untuk melatih model data pembelajaran

agar model sesuai dengan jenis parameter yang digunakan.

b. Data Uji

Data uji merupakan data yang digunakan dalam memberikan evaluasi melalui hasil model yang akan dilatih. Ketika data uji di *input* ke sistem, akan menjalani proses mengekstrak fitur dan membagi klasifikasi sesuai dengan model. Sistem akan mengeluarkan *output* klasifikasi oleh sistem yaitu negatif ataupun positif.

5. Klasifikasi

Klasifikasi data merupakan proses pembagian kategori pada pengelompokan data [17].

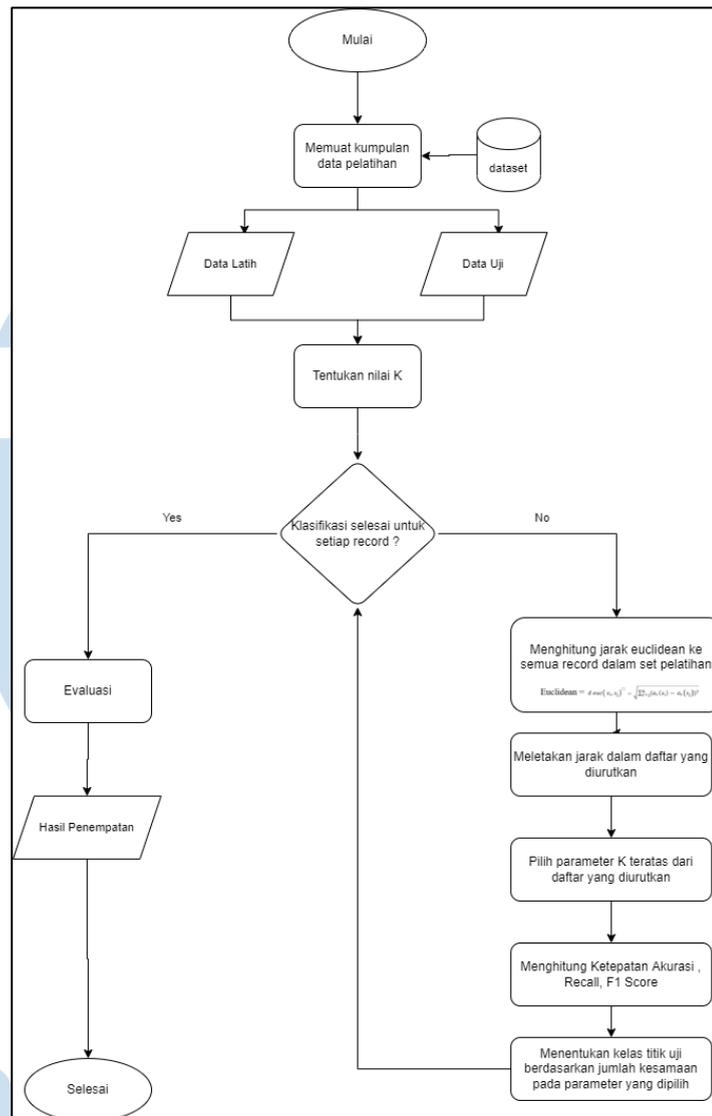
4. Model

Setelah model berhasil dibuat maka langkah selanjutnya adalah untuk mengevaluasi performa dengan akurasi dan presisi.

5. Evaluasi Sesuai

Setelah model dan evaluasi berjalan maka akan dilakukan proses pengecekan kembali. Apabila evaluasi tersebut tidak sesuai maka akan dilakukan pengulangan kembali dari membuat model dengan KNN, namun apabila evaluasi yang telah dilakukan sesuai maka sistem akan selesai.

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A



Gambar 3.3 Flowchart Implementasi Sistem

Pada Gambar 3.3 menunjukkan *flowchart* proses yang dilakukan untuk mendapatkan hasil pemodelan data yang diinginkan. Langkah pertama yang dilakukan yaitu memasukkan dataset yang sudah dibagi kedalam data latih dan data uji. Setelah itu pada tahap selanjutnya menghitung jarak terdekat pada semua *record* dalam data latih menggunakan rumus jarak Euclidean. Selanjutnya memilih parameter K yang akan dilakukan pada tahap uji data untuk menghitung nilai ketepatan akurasi, presisi, *recall* dan *F1 Score*. Dan pada tahap terakhir melakukan evaluasi pengujian dengan *cross validation* sebanyak 20 kali agar hasil data uji lebih optimal.