

BAB III METODOLOGI PENELITIAN

3.1 Subjek dan Objek Penelitian

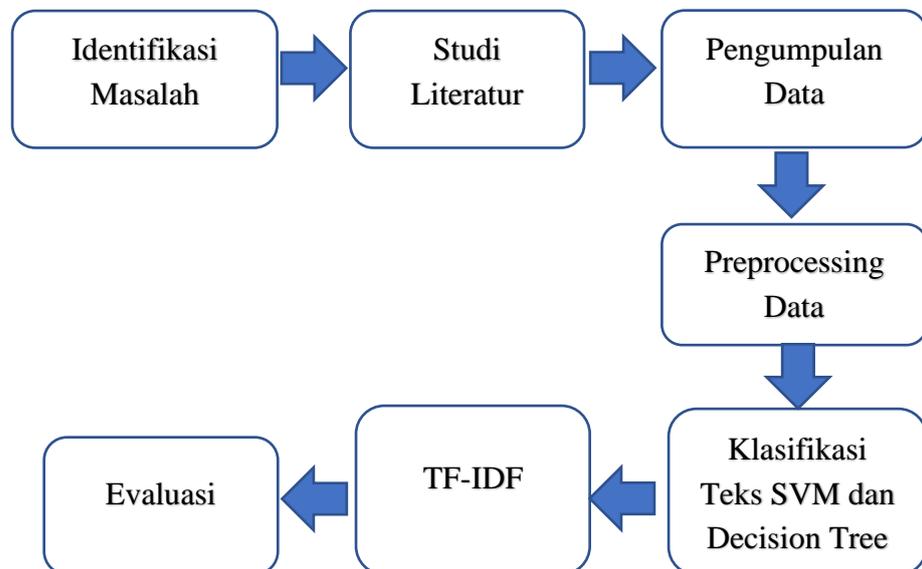
Subjek dalam penelitian ini adalah Institut Teknologi Telkom Purwokerto, sedangkan objek penelitiannya adalah Mahasiswa Institut Teknologi Telkom Purwokerto.

3.2 Alat dan Bahan Penelitian

Alat dan bahan yang akan digunakan pada penelitian adalah :

1. Laptop Acer dengan RAM minimal 4 GB, Processor ryzen 5, sistem operasi Windows 10
2. Browser Microsoft Edge
3. Data komentar survey kepuasan mahasiswa sebanyak 415 data
4. Jupyter Notebook

3.3 Proses Penelitian



Gambar 3. 1 Alur penelitian

a. Identifikasi Masalah

Identifikasi masalah proses untuk peneliti melakukan pendefinisian terhadap masalah yang akan diangkat dengan cara menentukan judul, latar belakang masalah tersebut, tujuan dan manfaat penelitian, oleh karena itu identifikasi masalah menjadi langkah awal penelitian yang sangat penting.

b. Studi Literatur

Studi literatur dimanfaatkan penulis untuk mencari referensi yang akan menjadi landasan dalam melakukan penelitian ini, dengan cara mencari informasi terkait penelitian yang akan dibahas di jurnal-jurnal yang telah ada, buku maupun internet.

c. Pengumpulan Data

Pengumpulan data merupakan proses untuk mengumpulkan dataset, penulis akan menggunakan data hasil survey kepuasan mahasiswa terhadap perkuliahan daring dari kaur BPP Institut Teknologi Telkom Purwokerto sebagai dataset nya, data yang digunakan berupa data komentar saja.

d. Preprocessing Data

Tahap ini merupakan tahap pembersihan data set sebelum mengklasifikasikan datanya, dimana data yang telah diperoleh dari hasil survey kepuasan mahasiswa terhadap perkuliahan daring diproses untuk menghasilkan data yang diperlukan saja dan menghilangkan data yang tidak diperlukan. Teknik memproses data tersebut adalah sebagai berikut :

1. *Case Folding*, yaitu data kuisisioner yang ada akan dilakukan perubahan menjadi sama *lowercase* (membuat semua huruf menjadi kecil).

2. *Tokenization*, yaitu kalimat yang ada akan dipecah menjadi beberapa kata dan juga menghilangkan tanda baca yang ada.

3. *labelling* dilakukan secara manual, dengan cara membagi kelas data kedalam dua kelas yaitu kelas negatif dan kelas positif, data yang dilabelkan secara manual ini berfungsi sebagai dataset untuk proses klasifikasi dengan metode SVM dan Decision Tree.

4. *Filtering*, yaitu proses menghilangkan kata yang tidak memiliki kontribusi besar atau tidak memiliki pengaruh.

5. *Stemming*, yaitu proses untuk mengubah kalimat menjadi kata kata dasar dengan cara menghilangkan semua imbuhan [4].

e. TF-IDF

Pada proses ini dilakukan pembobotan kata pada data komentar yang ada, komentar yang merupakan teks akan berubah menjadi data numerik agar bisa diproses oleh sistem.

f. Klasifikasi teks

Pada tahap klasifikasi ini menghasilkan sebuah ulasan yang terbagi menjadi dua kelas, yaitu ulasan positif dan ulasan negatif dengan menggunakan algoritme SVM dan Decision Tree dengan menggunakan empat skenario yaitu 90:10, 80:20, 70:30 dan 60:40 [4].

Tabel 3. 1 Jumlah data untuk setiap skenario

Skenario	Data Training	Data testing
90:10	307	35
80:20	273	69
70:30	239	103
60:40	205	137

g. Evaluasi

Setelah klasifikasi telah selesai dilakukan maka selanjutnya yang akan dilakukan adalah evaluasi kedua algoritma tersebut dengan menggunakan metode Confusion Matrix, dimana metode ini melakukan perhitungan dengan hasil berupa empat keluaran yaitu *recall*, *precision*, *accuracy* dan *f1-score* [13].