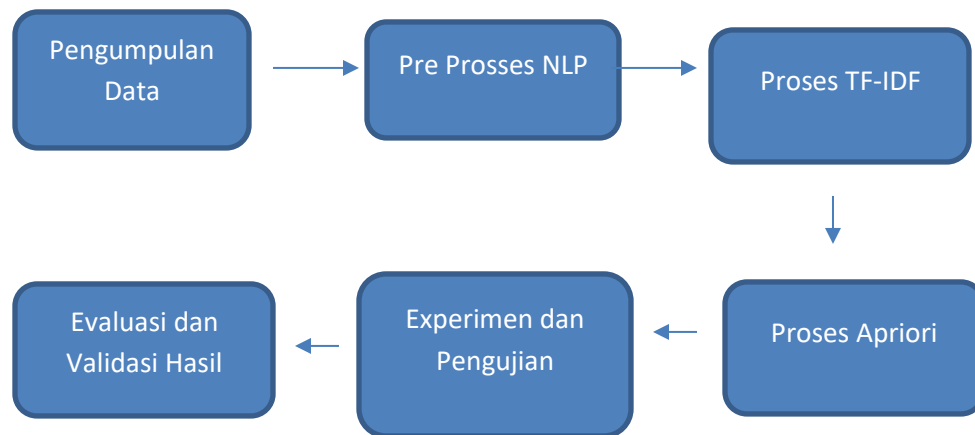


## BAB III

### METODE PENELITIAN

#### A. Desain Penelitian

Metode penelitian yang digunakan adalah metode eksperimen. Metode eksperimen adalah metode cara penyajian sebuah data melalui proses percobaan-percobaan untuk membuktikan sendiri suatu pernyataan atau hipotesis tertentu dalam sebuah penelitian. Keuntungan utama penelitian dengan penggunaan metode eksperimen adalah adanya kendali ditangan peneliti dan ketepatan sebuah logika yang terkandung di dalam metode tersebut (Setyanto et al, 2006). Metode ini juga sering digunakan karena dapat dibuat perencanaan yang sistematis sehingga metode eksperimen adalah metode yang paling kuat dalam sebuah penelitian. Berikut adalah tahapan metode eksperimen :



**Gambar 3.1 Metode yang Diusulkan**

## B. Pengumpulan Data

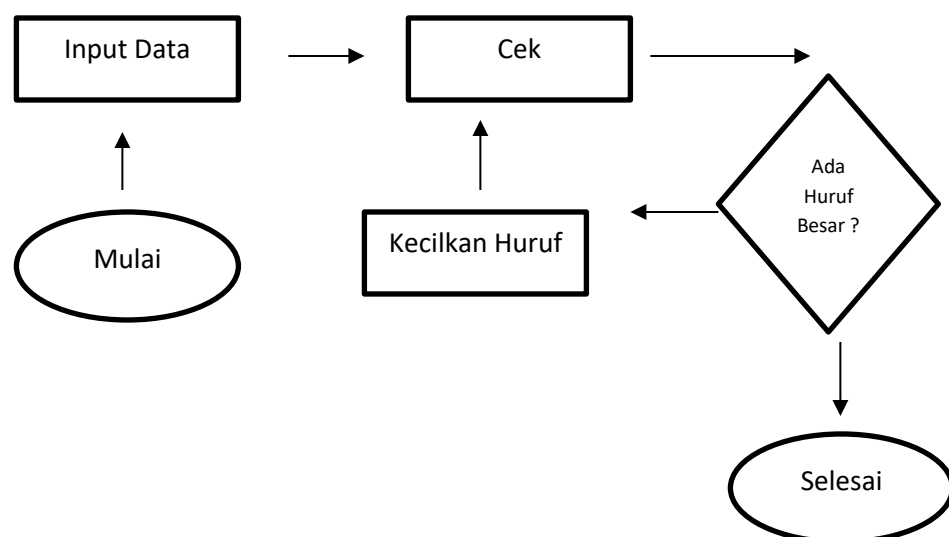
Data yang dikumpulkan terdiri dari 2 buah jenis data yakni data uji dan data latih. Data latih yakni data yang berisikan kumpulan berita *hoax* yang nantinya akan digunakan untuk proses pengindex an kata guna pembuatan *library*. Data uji yakni data yang berisikan berita baik *hoax* maupun real yang nantinya digunakan untuk menentukan tingkat akurasi sistem. Data yang diambil adalah data berita *hoax* dari situs *turnbackhoax.id* dan juga data *hoax* dari broadcast whatsapp. Sedangkan berita yang nantinya dijadikan data uji berasal dari *detik.com* dan sebagian berasal dari *turnbackhoax.id*.

## C. Pre Prossesing Menggunakan NLP (Neural Language Processing)

Data berita yang didapat dari situs *turnbackhoax.id* ataupun dari *detik.com* adalah sebuah data mentah yang belum dapat diolah. Agar data tersebut dapat diolah, maka dilakukan pre prossesing dengan menggunakan NLP (*Neural Language Processing*). Pada tahap pre prossesing ini memiliki beberapa *step* yakni :

### 1. Case Folding

Perubahan huruf besar menjadi huruf kecil pada berita.



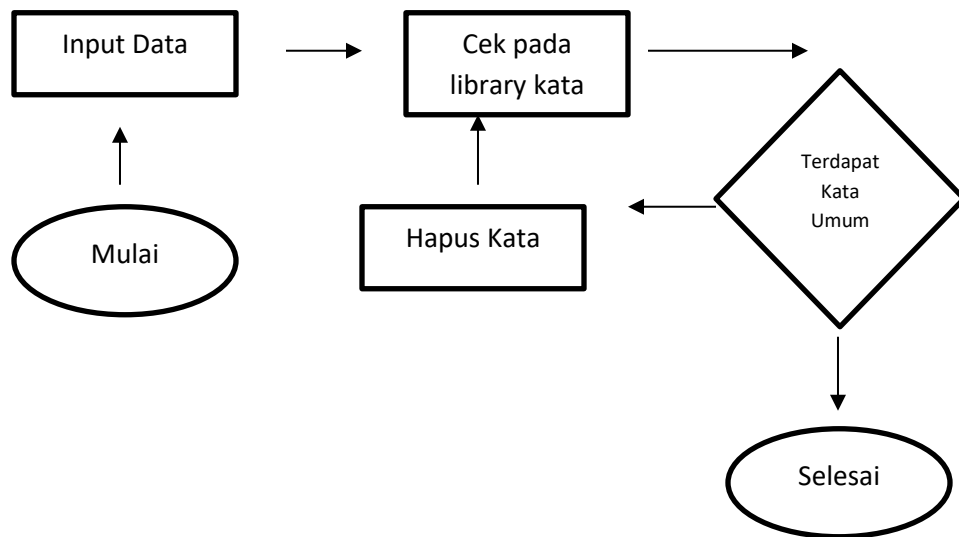
**Gambar 3.1 Flowchart Case Folding**

## 2. Tokenizing

Proses pemisahan text menjadi per kata sehingga tiap kata dapat diolah dengan mudah. Proses pemisahan kata dilakukan menggunakan fungsi dalam python.

## 3. *Stoward Removal*

Stopword merupakan kata umum yang digunakan / kata yang tidak penting untuk dianalisi / diproses. Seperti contoh kata : dan, atau, akan tetapi, jika. Tujuan dari penghapusan kata ini yakni mengurangi pembengkakan jumlah index yang digunakan.



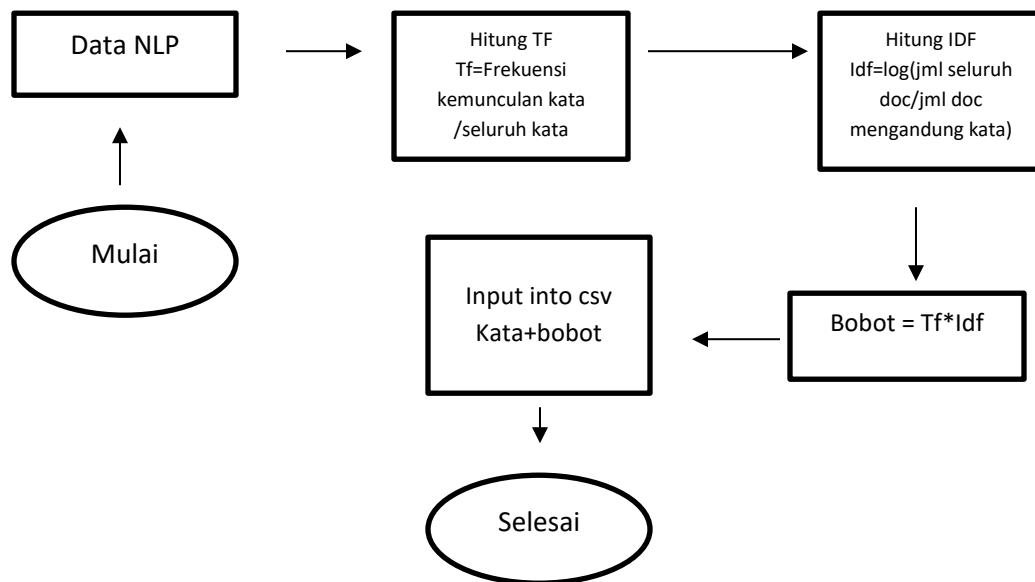
**Gambar 3.2 *Flowchart Stoward Removal***

#### 4 Stemming

Sebuah penghapusan kata yang memiliki awalan / akhiran sehingga didapat kata dasar. Guna dari *stemming* ini yakni meminimalisir persamaan kata yang memiliki perbedaan awalan / akhiran sehingga dapat memperkecil jumlah indeks yang akan diproses. Proses stemming ini menggunakan fungsi *stemmer factory* yang terdapat pada python dengan *import package* sastrawi

#### D. Processing menggunakan Tf-Idf

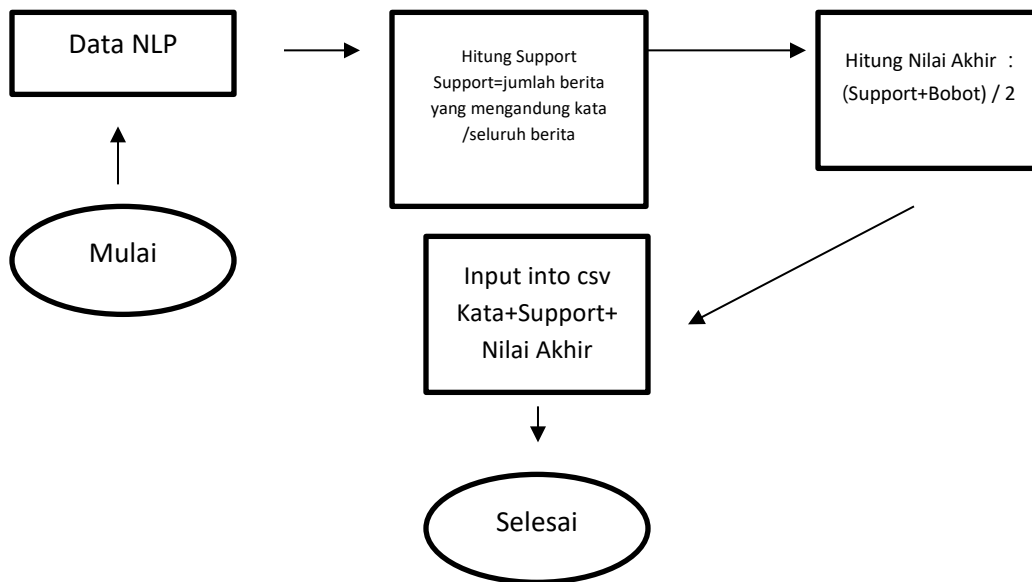
Setelah dilakukan *pre processing* dengan menggunakan NLP, maka langkah selanjutnya yakni mencari bobot dari sebuah kata dengan menggunakan *Term Frequent* dan *Inverse Document Frequent*. Alur sistem dari program Tf-Idf ini digambarkan pada *flowchart* berikut :



**Gambar 3.3 Flowchart Tf-Idf**

### E. *Processing menggunakan Apriori*

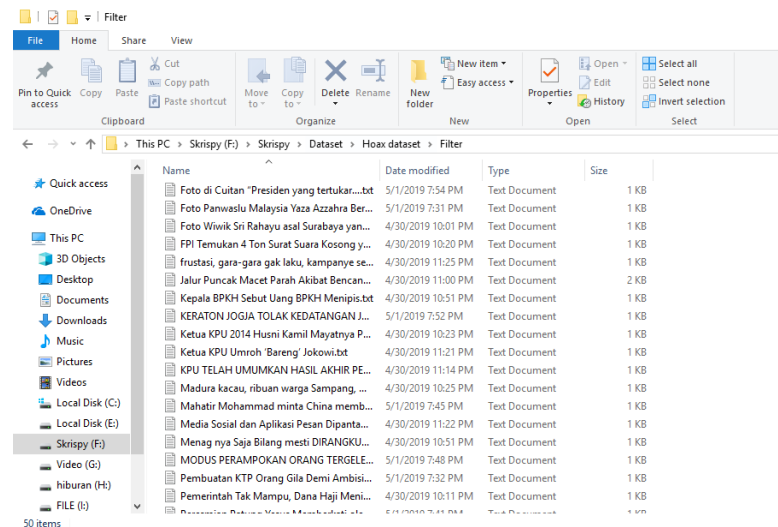
Tahap selanjutnya yakni prosesing dengan menggunakan metode apriori. Dalam metode ini diambil sebuah nilai support dari perbandingan antara jumlah berita yang mengandung kata dengan jumlah total berita. Yang selanjutnya akan dimasukkan kedalam data *index library* kata yang tersimpan beserta bobot. Selanjutnya Data akan dicari rata-rata dari nilai bobot dan nilai support sehingga didapat nilai akhir. Proses alur dari sistem digambarkan pada *flowchart* berikut :



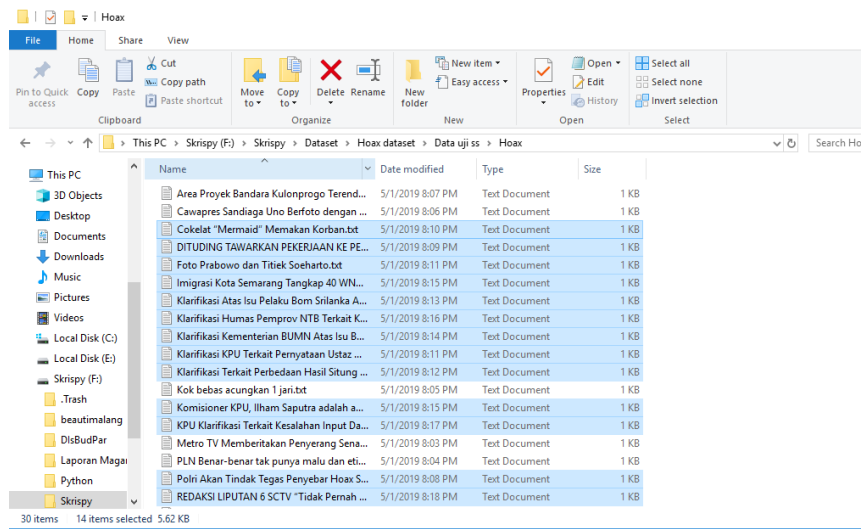
**Gambar 3.4 *Flowchart Apriori***

### F. *Eksperimen dan Pengujian*

Penelitian ini akan menggunakan data latih sebanyak 50 buah berita *hoax* yang dapat di lihat pada Lampiran 1. Dan 30 buah data uji dengan rincian 16 buah berita *hoax* dan 14 buah berita real yang dapat dilihat pada Lampiran 2. Dengan data uji tersebut dilakukan pengujian sistem untuk memastikan tidak terjadi error pada saat proses pengklasifikasian. Berikut adalah beberapa berita yang akan digunakan sebagai data uji dan data latih :



**Gambar 3.5 Data Latih**



**Gambar 3.6 Data Uji**

#### **G. Evaluasi dan Validasi Hasil**

Evaluasi dilakukan dengan memasukkan hasil dari data uji yang telah diuji pada tahap sebelumnya. Digunakan rumus *precision and recall* dan *accuracy* untuk menentukan tingkat keakuratan sistem.