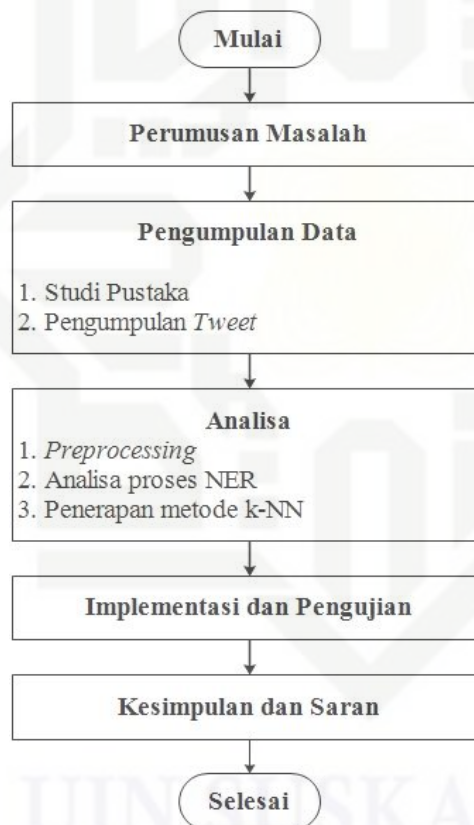


## BAB III

### METODOLOGI PENELITIAN

Metodologi penelitian merupakan langkah-langkah sistematis yang akan dilakukan dalam sebuah penelitian untuk menyelesaikan masalah dalam penelitian tersebut, supaya penelitian yang dilakukan akan lebih terarah dan memperoleh hasil maksimal sesuai perencanaan. Berikut adalah tahapan penelitian yang akan dilakukan :



**Gambar 3.1 Metodologi Penelitian**

#### 3.1 Perumusan Masalah

Tahapan ini merupakan tahapan awal dari metodologi penelitian. Pada tahap ini dilakukan suatu perumusan masalah yaitu bagaimana *Named Entity Recognition*

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

dengan menggunakan metode *k-Nearest Neighbor* untuk mengenali nama entitas pada *tweet* iklan.

## 3.2 Pengumpulan Data

Tahap pengumpulan data dilakukan untuk mengumpulkan informasi. Data yang dikumpulkan pada penelitian ini adalah sebagai berikut:

### 3.2.1 Studi Pustaka

Studi pustaka untuk mencari referensi yang sesuai dengan permasalahan yang diteliti. Referensi yang ditemukan dapat berupa paper, jurnal, buku dan beberapa situs di internet dengan topik terkait. Studi pustaka bertujuan agar penelitian yang dilakukan memiliki landasan atau dasar penelitian sehingga memperkuat permasalahan.

### 3.2.2 Pengumpulan *Tweet*

Pengumpulan data *tweet* dilakukan dengan menggunakan Twitter API (*Application Programming Interface*). Data *tweet* yang dikumpulkan merupakan *tweet* yang termasuk iklan dan berjumlah 1000 *tweet*.

## 3.3 Analisa

Pada tahapan ini dilakukan analisa untuk *Named Entity Recognition* (NER) pada *tweet* iklan menggunakan metode *k-Nearest Neighbor* (k-NN). Tahap analisa yang akan dilakukan yaitu:

### 3.3.1 *Preprocessing*

*Preprocessing* adalah tahapan untuk mempersiapkan teks menjadi data yang akan diolah di tahapan berikutnya. Inputan awal pada proses ini adalah berupa *tweet* iklan. *Preprocessing* pada penelitian ini terdiri dari beberapa tahapan, yaitu: *cleaning*, *case folding*, *tokenizing* dan normalisasi. *Cleaning* adalah proses pengapusan tanda baca. *Case folding* adalah mengubah semua huruf dalam dokumen menjadi huruf kecil. *Tokenizing* adalah proses menghasilkan kata-kata yang berdiri sendiri dan Normalisasi adalah mengubah kata tidak baku menjadi baku.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

### 3.3.2 Analisa proses NER

Setelah tahapan *preprocessing*, kemudian dilakukan analisa proses NER. Pada proses pengenalan nama entitas (NER). Tahapan yang dilakukan adalah sebagai berikut:

1. POS *Tagging* dilakukan untuk menandai kelas kata dalam teks (misalnya kata benda, kata kerja, kata sifat).
2. *Chunking* dilakukan untuk mengidentifikasi kata yang merupakan frasa dengan kombinasi dari kelas kata.
3. Kemudian dilakukan *encoding* BIO (*Beginning, Inside, Outside*) untuk penandaan nama entitas.

### 3.3.3 Penerapan Metode k-NN

Setelah proses NER, kemudian dilakukan penerapan metode k-NN untuk klasifikasi nama entitas. Pada tahap ini, dilakukan pembobotan kata menggunakan TF-IDF kemudian dilakukan penerapan metode k-NN. Pengenalan nama entitas yang diklasifikasi dikelompokkan menjadi empat kategori yaitu nama penjual, nama produk, alamat dan kontak.

## 3.4 Perancangan

Pada tahap ini dilakukan perancangan sistem yaitu dengan membuat *flowchart* untuk mengetahui gambaran seluruh proses terhadap masukan dan keluaran pada sistem dan melakukan perancangan basis data yang mencakup perancangan tabel.

## 3.5 Implementasi dan Pengujian

Tahap implementasi dan pengujian merupakan tahapan yang dilakukan setelah tahapan analisa dan perancangan telas selesai. Pada tahap ini menjelaskan tentang penerapan sistem yang telah dianalisa dan dirancang sebelumnya.

### 3.5.1 Implementasi

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Pada tahapan implementasi dilakukan setelah analisa dan perancangan sistem selesai. Sistem yang telah dirancang diimplementasikan dalam bentuk *source code*. Untuk dapat mengimplementasikan sistem ini, maka dibutuhkan perangkat pendukung, seperti perangkat keras (*hardware*) dan perangkat lunak (*software*).

1. Perangkat keras (*Hardware*), antara lain:

<i>Processor</i>	: Intel(R) Core 2
<i>Memory (RAM)</i>	: 2 GB
<i>System Type</i>	: 32 bit Operating Sistem
<i>Hard disk</i>	: 500 GB

2. Perangkat lunak (*Software*), antara lain:

<i>Operating System</i>	: Windows 10
Bahasa Pemrograman	: PHP
<i>Tools</i>	: <i>Sublime Text 3</i>

### 3.5.2 Pengujian

Pada tahap pengujian, sistem akan diuji untuk mengetahui hasil analisa dan perancangan pada sistem serta akurasi dari proses sistem. Untuk mengetahui validitas sistem, maka akan dilakukan pengujian terhadap hasil yang diperoleh dari sistem. Tahapan pengujian dilakukan dengan 2 cara yaitu pengujian tingkah laku sistem yang telah dirancang menggunakan *white box* dan *confusion matrix* untuk melakukan perhitungan akurasi.

### 3.6 Kesimpulan dan Saran

Pada tahapan ini memberikan kesimpulan dari penelitian yang dilakukan yaitu *Named Entity Recognition* pada *tweet* iklan menggunakan metode *k-Nearest Neighbor*. Selanjutnya memberikan saran terhadap topik yang diteliti sehingga dapat menghasilkan penelitian baru yang lebih baik lagi.