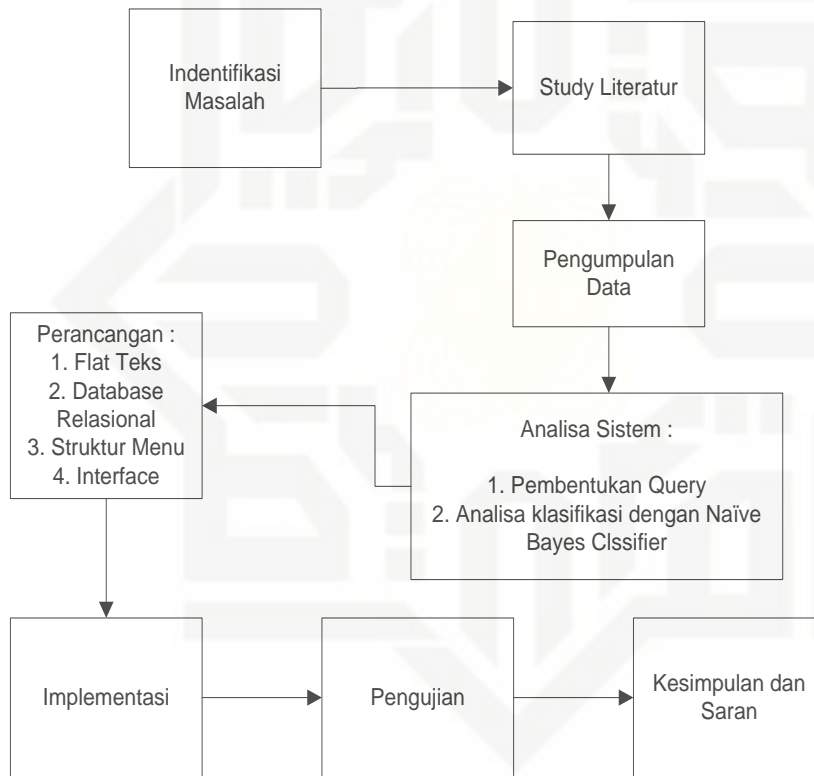


BAB III

METODOLOGI PENELITIAN

Penelitian yang dilakukan memiliki tahapan-tahapan dalam pelaksanaannya, sehingga di dalam pelaksanaannya tersusun secara sistematis dan dapat mencapai tujuan yang diharapkan. Maka dari itu disusunlah tahapan-tahapan metodologi dalam penelitian ini yang akan dijelaskan pada Gambar 3.1 dibawah ini :



Gambar 3.1 Tahapan Penelitian

Berdasarkan Gambar 3.1 terdapat beberapa tahapan dalam penelitian ini. Akan dijelaskan secara rinci tahapan pada penulisan dibawah ini.



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.

3.1 Identifikasi Masalah

Tahap ini merupakan langkah awal dari metodologi penelitian dengan mencari dan mempelajari masalah berdasarkan topik yang akan diteliti. Kemudian menentukan ruang lingkup, latar belakang dan bagaimana mencari solusi dalam masalah tersebut. Dalam penelitian ini berdasarkan latar belakang yang telah dikemukakan, maka dirumuskan permasalahan yaitu bagaimana membangun sebuah aplikasi untuk mengkalasifikasikan topik berita dengan menggunakan metode *Naive Bayes*.

3.2 Study Pustaka / Literatur

Tahapan ini merupakan tahapan dimana dilakukan pengumpulan data dan informasi yang berhubungan dengan penelitian ini untuk lebih mengetahui tentang masalah yang akan diteliti. Studi pustaka yang dilakukan dengan tujuan untuk mencari teori-teori yang dapat dijadikan acuan dan mendukung penelitian dengan membaca buku-buku referensi, jurnal, paper, situs internet dan penelitian sebelumnya yang berhubungan.

Pengumpulan data berasal dari situs berita *online Detik.us* berupa Dokumen, adapun jumlah data yang digunakan berjumlah 150 data terdiri dari 50 data diambil dari berita Hiburan, 50 data diambil dari berita Olahraga dan 50 data diambil dari berita Kesehatan pada *Detik.us*.

3.3 Pengumpulan Data

Tahap ini merupakan tahap dimana data yang dibutuhkan dalam penelitian ini dikumpulkan untuk dapat diolah dan digunakan dalam pembuatan sistem. Data-data yang dikumpulkan adalah sebagai berikut:

a. Topik Berita

Adapun sumber data, metode pengumpulan data dan jumlah data adalah sebagai berikut :

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.

1. Sumber berita berasal dari *Detik.us*.
2. Metode pengumpulan data dilakukan secara manual dari bulan Agustus 2016 sampai dengan Januari 2017.
3. Jumlah data yang dikumpulkan berjumlah 150 data dengan format *.txt, setelah terkumpul kemudian membagi data menjadi perbandingan data latih dan data uji sebagai berikut: 70:30, 60:40, dan 50:50.

b. Kategori Topik Berita

Penentuan kategori topik Berita dilakukan oleh administrator pada bagian Kategori, Namun situs-situs berita *online* tidak semua berita yang berhubungan dengan kata kunci Olahraga, Kesehatan dan Hiburan merupakan berita yang menyatakan termasuk berita tersebut. Hasil diperoleh dari kategori yaitu “olahraga”, “Kesehatan” dan “Hiburan”, selanjutnya memberikan kategori pada tiap dokumen berdasarkan kategori.

3.4 Analisa Sistem

Setelah melakukan studi pustaka, selanjutnya melakukan analisa yang bertujuan untuk mengetahui, memperoleh gambaran yang jelas mengenai masalah, serta menganalisa data yang dibutuhkan pada penelitian. Pada tahapan ini akan dijelaskan secara rinci tentang proses sistem yang akan dibangun. Dalam penelitian ini akan dilakukan klasifikasi terhadap data-data yang dikumpulkan untuk mengklasifikasi topik berita menggunakan metode *Naive Bayes*. Berikut adalah rincian dari analisa yang dilakukan :

Analisa Sistem pada *query* menggunakan Algoritma *Naive Bayes*. Tahapan ini terdiri dari beberapa langkah, yakni

- a. *Preprocessing query*. Tahapan yang dilakukan adalah :

Case folding, menghapus karakter selain UTF-8, link URL dan dokumen yang berulang, *tokenization*, melakukan *linguistic preprocessing*, normalisasi dan Pembobotan kata

1. Dokumen yang didapat dari kumpulan berita sebanyak 150 data berhubungan dengan topik berita yang ada sebagai data latihnya 120 data,

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.

dan data uji 30 data, setiap dokumen secara manual. Proses selanjutnya akan dilakukan melalui sistem.

2. *Case folding* : tahapan ini merubah semua huruf kapital dengan huruf kecil.
3. Menghapus karakter selain UTF-8, link URL dan dokumen yang berulang
4. Pemisahan rangkaian kata (tokenization).
5. Melakukan *linguistic preprocessing*: *stopword* dan *stemming*
6. Normalisasi : tahapan identifikasi kata slang dan penulisan kata berlebihan kemudian di ganti dengan kata kamus KBBI
7. *Indexing*.
8. Pembobotan kata.
- b. Klasifikasi *query* dengan algoritma *Naive bayes*. Tahapan yang dilakukan
 1. Menghitung probabilitas pada setiap kelas topik berita
 2. Cari probabilitas nilai tertinggi
 3. Menghitung akurasi

3.5 Perancangan

Perancangan berarti metode yang khusus digunakan untuk merancang sistem yang telah dianalisa dengan tujuan untuk memberikan kemudahan dan menyederhanakan suatu proses atau jalannya aliran data, perancangan terhadap model, dan merancang rancang bangun sistem. Adapun rancangan utama sistem yaitu:

1. Perancangan Flat Teks

Merancang penyimpanan data dalam flat teks (*plain text*)
2. Perancangan Database Relasional

Merancang penyimpanan data dalam konseptual data model
3. Perancangan struktur menu

Merancang menu-menu pada aplikasi sesuai dengan fungsi dan kegunaan masing-masing.
4. Perancangan antarmuka (*interface*)


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

Merancang antarmuka (*interface*) sistem yang akan membuat interaksi antara pengguna (*user*) dengan sistem.

3.6 Implementasi

Pada proses implementasi ini akan dilakukan pembuatan modul yang telah dirancang dan dianalisa selanjutnya diimplementasikan pada bahasa pemrograman dan dilakukan pengujian untuk mengetahui tingkat keberhasilan aplikasi yang telah ada. Berikut adalah spesifikasi lingkungan implementasi perangkat keras dan perangkat lunak :

1. Perangkat keras

Processor	: Intel® Core (TM)2 Duo CPU 2.00GHz
Memori (RAM)	: 2,00 GB
Harddisk	: 300 GB

2. Perangkat Lunak

Sistem Operasi	: <i>Windows Vista Basic</i>
<i>Tools</i> perancangan	: Notepad++ v6.6.8
<i>Web Browser</i>	: Google Chrome (<i>Version 37</i>) / Mozilla 5.0
Bahasa pemrograman	: PHP
DBMS	: MySQL
Perangkat pendukung	: XAMPP 1.8.2

3.7 Pengujian

Pengujian merupakan tahapan dimana sistem akan dijalankan. Tahap pengujian diperlukan sebagai ukuran bahwa sistem dapat dijalankan sesuai dengan tujuan. Pengujian perangkat lunak, pengujian yang dilakukan adalah pengujian *blackbox* yaitu dengan menguji apakah perangkat lunak yang dibangun sesuai dengan rancangan dan keluaran yang diharapkan. Pengujian sistem klasifikasi berita ini dilakukan dengan model *confusion matrix* yaitu, sebuah matriks prediksi yang akan



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

dibandingkan dengan kelas asli dari data inputan. Dalam penelitian ini, dibandingkan hasil dari keluaran sistem dengan hasil yang diharapkan dengan permodelan *confusion matrix*

3.8 Kesimpulan dan Saran

Tahapan ini berisikan tentang intisari penelitian dan hasil yang didapatkan menggunakan metode *Naive Bayes* dalam melakukan klasifikasi terhadap topik berita, serta berisikan hal yang disimpulkan dan disarankan penulis bagi pembaca untuk melakukan pengembangan terhadap penelitian ini kedepannya.

