

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang Masalah

Kaum muslimin dalam kehidupan sehari-harinya tidak dapat terlepas dari hukum Islam, salah satunya yaitu dalam konteks taharah. Secara bahasa, kata taharah berasal dari kata “*thahara-yathuru-thaharatan*” yang berarti bersuci [1]. Hukum *taharah* adalah wajib, dan tidak sah apabila dalam pengerjaannya tidak sesuai syariat Al-Qur’an dan As-Sunnah. Diantara ayat-ayat Al-Qur’an yang menjelaskan tentang kewajiban taharah adalah :

يَا أَيُّهَا الَّذِينَ آمَنُوا إِذَا قُمْتُمْ إِلَى الصَّلَاةِ فَاغْسِلُوا وُجُوهَكُمْ وَأَيْدِيَكُمْ إِلَى الْمَرَافِقِ وَامْسَحُوا بِرُءُوسِكُمْ وَأَرْجُلَكُمْ إِلَى الْكَعْبَيْنِ  
وَإِنْ كُنْتُمْ جُنُبًا فَاطَّهَّرُوا وَإِنْ كُنْتُمْ مَرْضَىٰ أَوْ عَلَىٰ سَفَرٍ أَوْ جَاءَ أَحَدٌ مِنْكُم مِّنَ الْغَائِطِ أَوْ لَامَسْتُمُ النِّسَاءَ فَلَمْ تَجِدُوا مَاءً  
فَتَيَمَّمُوا صَعِيدًا طَيِّبًا فَامْسَحُوا بِوُجُوهِكُمْ وَأَيْدِيكُمْ مِنْهُ مَا يُرِيدُ اللَّهُ لِيَجْعَلَ عَلَيْكُمْ مِنْ حَرَجٍ وَلَٰكِنْ يُرِيدُ لِيُطَهَّرَكُمْ وَلِيَنبِتَ  
نِعْمَتَهُ عَلَيْكُمْ لَعَلَّكُمْ تَشْكُرُونَ

UNIVERSITAS ISLAM NEGERI  
SUNAN GUNUNG DJATI  
BANDUNG

“Hai orang-orang yang beriman, apabila kalian hendak mengerjakan salat, maka basuhlah muka kalian dan tangan kalian sampai dengan siku, dan sapulah kepala kalian dan (basuh) kaki kalian sampai dengan kedua mata kaki; dan jika kalian junub, maka mandilah; dan jika kalian sakit atau dalam perjalanan atau kembali dari tempat buang air (kakus) atau menyentuh perempuan, lalu kalian tidak memperoleh air, maka bertayammumlah dengan tanah yang baik (bersih); sapulah muka kalian dan tangan kalian dengan tanah itu. Allah tidak hendak menyulitkan kalian, tetapi Dia hendak membersihkan kalian dan

*menyempurnakan nikmat-Nya bagi kalian, supaya kalian bersyukur. (QS. Al Maidah : 6)''.*

Terdapat banyak media yang memberikan informasi tentang taharah seperti buku-buku formal, video tutorial, artikel dan lain sebagainya. Namun saat ini, sudah hadir teknologi-teknologi yang mendukung komputer bertindak lebih cerdas, diantaranya *Artificial Intelligence* (Kecerdasan Buatan). Kita dapat memanfaatkan teknologi AI tersebut untuk memberikan informasi kepada muslim tentang taharah.

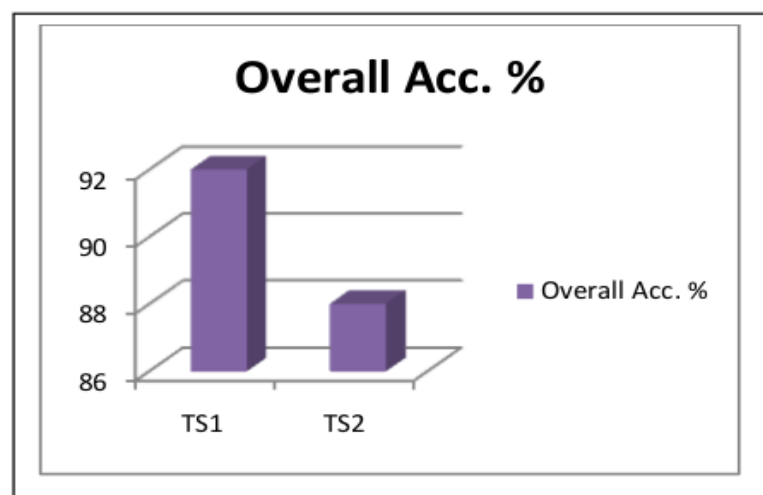
Menurut Kusumadewi (2003), *Natural Language Processing* atau NLP merupakan bidang *Artificial Intelligence* yang mempelajari komunikasi antara komputer dan manusia melalui *natural language* atau bahasa alami. Salah satu bentuk aplikasi NLP adalah *chatbot*. Kemampuan *chatbot* untuk menjawab pertanyaan (*question answering*) dengan bahasa alami dapat dimanfaatkan sebagai *conversational search engine*, yaitu mesin pencari dengan bahasa yang konversasional. Pada umumnya, proses *question answering* pada *chatbot* memiliki alur seperti berikut ini :



**Gambar 1.1** Alur kerja *chatbot* pada *question answering* [2]

Untuk menganalisa pertanyaan dari *user*, *chatbot* mengidentifikasi *intent* dan *entity* pertanyaan terlebih dahulu. *Intent* pada *chatbot* merupakan maksud yang ingin diketahui oleh *user*. Sedangkan *entity* mewakili konteks spesifik dari *intent*. Sebagai contoh, saat *user* menginputkan pertanyaan “Dimana letak bangunan A?”, *intent* atau yang dimaksud oleh *user* adalah “lokasi” sedangkan *entity* atau konteks spesifiknya adalah “lokasi bangunan A”.

*Machine Learning* (ML) yaitu pembelajaran mesin merupakan pendekatan AI yang banyak digunakan untuk menirukan perilaku manusia dalam menyelesaikan masalah atau melakukan otomatisasi. Ciri khas dari *Machine Learning* adalah adanya proses *data-training*, karena *machine learning* belajar dari data sebagai *knowledge*-nya. Dalam penelitian yang dilakukan oleh Nanda et al. (2016) dinyatakan bahwa pada sistem *question answering*, konsep *similarity* dengan pendekatan *machine learning* (TS1) lebih efektif dari pada konsep *equal* (TS2), yang dibuktikan dengan hasil akurasi ketepatan jawaban diatas 90%. Hasil akurasi dapat dilihat di gambar 1.2



**Gambar 1.2** Accuracy dari TS1 dan TS2 [3]

Klasifikasi merupakan metode dalam ML yang digunakan untuk mengklasifikasi objek data sebagai salah satu kategori (kelas) yang telah didefinisikan sebelumnya[4]. Klasifikasi digunakan oleh mesin untuk memilah objek berdasarkan ciri tertentu, sebagaimana manusia mencoba membedakan benda satu dengan yang lain. Salah satu algoritma klasifikasi yang umum digunakan adalah *Naive Bayes*.

Algoritma *Naive Bayes* menggunakan metode probabilitas dan statistik yang ditemukan oleh ilmuwan bernama Thomas Bayes di abad ke-18. Algoritma *Naive Bayes* bekerja lebih maksimal dibanding dengan algoritma lainnya. Hal ini dibuktikan dalam penelitian yang dilakukan oleh Xhemali et al. (2009), pada penelitian tersebut dikatakan bahwa *Naive Bayes Classifier* memiliki akurasi yang lebih baik dibanding *model classifier* lainnya. Berikut tabel hasil komparasi algoritma *Naive Bayes* dengan *Decision Tree* :

**Tabel 1.1** Accuracy dari *Naive Bayes* dan *Decision Tree*[5]

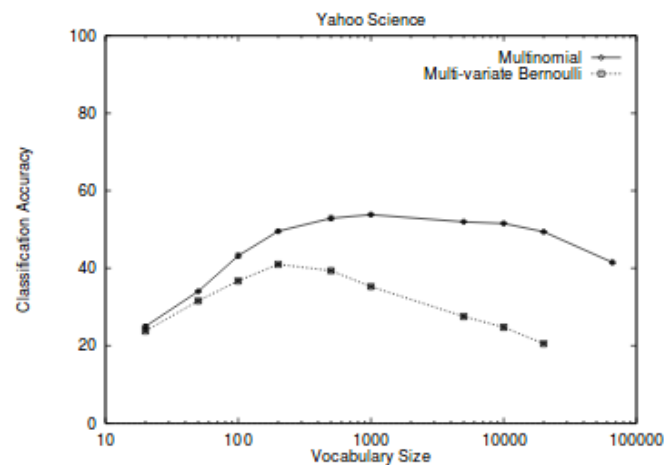
| <i>Classifier</i>    | <i>Accuracy</i> | <i>Precision</i> | <i>Recall</i> | <i>F-Measure</i> |
|----------------------|-----------------|------------------|---------------|------------------|
| <i>NB Classifier</i> | <b>95.20%</b>   | 99.37%           | 95.23%        | 97.26%           |
| <i>DT Classifier</i> | 94.84%          | 98.31%           | 95.90%        | 97.09%           |

Selain itu, pada penelitian Sudha dan Govindarajan (2016), dikatakan juga bahwa algoritma *Naive Bayes* meraih hasil akurasi lebih baik dibandingkan dengan algoritma *Support Vector Machine* yang dibuktikan pada tabel 1.2.

**Tabel 2.2** Accuracy *Naive Bayes* dan *Support Vector Machine* [6]

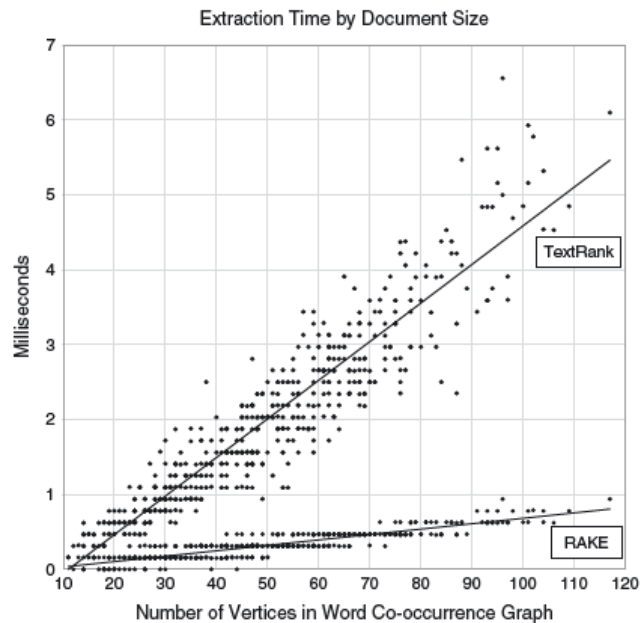
| <i>Dataset</i>           | <i>Algorithm</i>                    | <i>Accuracy</i> |
|--------------------------|-------------------------------------|-----------------|
| <i>Movie review data</i> | <i>Naive Bayes (NB)</i>             | <b>87.9%</b>    |
|                          | <i>Support Vector Machine (SVM)</i> | 86.5%           |

Terdapat beberapa model klasifikasi dari *Naive Bayes*, diantaranya *Multinomial Naive Bayes* dan *Multivariate Bernouli Naive Bayes*. McCallum et al. (1998) pada penelitiannya menyatakan bahwa klasifikasi *Multinomial Naive Bayes* dapat menangani kosakata dalam jumlah besar dengan cukup baik yang dibuktikan pada gambar 1.3. Sehingga klasifikasi *Multinomial Naive Bayes* dapat digunakan pada mesin *chatbot* untuk mengidentifikasi *intent* dari pertanyaan yang diinputkan.

**Gambar 1.3** Perbandingan model yang berbeda menggunakan dataset Yahoo[7]

Sedangkan untuk mengidentifikasi *entity*, mesin *chatbot* dapat menggunakan pendekatan ML lainnya, diantaranya dengan metode *keyword extraction* atau ekstraksi kata kunci. Berdasarkan penelitian yang dilakukan oleh

Stuart Rose (2010), metode *keyword extraction* RAKE bekerja lebih cepat dibandingkan *keyword extraction* lainnya (TextRank) yang dibuktikan dengan grafik pada gambar 1.4



**Gambar 1.4** Perbandingan kinerja algoritma RAKE dan TextRank [8]

Kemampuan *conversational search engine* pada *chatbot* juga seharusnya memungkinkan *back-and-forth communication* atau komunikasi bolak balik antara *user* dengan *bot*. *Back-and-forth communication* yang dimaksud adalah pertanyaan balik yang diberikan oleh *chatbot* untuk mendapatkan pertanyaan yang lebih spesifik dari user. Akan tetapi, dari hasil studi literatur yang dilakukan pada beberapa penelitian sebelumnya, *chatbot* yang dirancang belum mampu melakukan *back-and-forth communication*.

Berdasarkan pemaparan tersebut, peneliti tertarik untuk mengimplementasikan algoritma *Multinomial Naïve Bayes* sebagai *intent classifier* dan mengimplementasikan *Rapid Automatic Keywords Extraction*

sebagai *entity recognizer* pada *chatbot* tanya jawab *taharah*, yang juga dilengkapi oleh fitur *back-and-forth communication*.

## 1.2 Perumusan Masalah

Dari uraian latar belakang yang telah dijelaskan, maka dapat dirumuskan masalah sebagai berikut :

- a. Bagaimana membangun *intent classifier* pada mesin *chatbot* menggunakan *Multinomial Naïve Bayes*?
- b. Bagaimana membangun *entity recognizer* dan *back-and-forth communication system* pada mesin *chatbot* menggunakan *Rapid Automatic Keyword Extraction*?
- c. Bagaimana akurasi jawaban yang dihasilkan oleh aplikasi *chatbot* untuk menjawab pertanyaan *user*?

## 1.3 Tujuan Penelitian

Adapun tujuan dari membangun aplikasi *chatbot* hukum *taharah* ini adalah :

- a. Membangun *intent classifier* pada mesin *chatbot* menggunakan *Multinomial Naïve Bayes*.
- b. Membangun *entity recognizer* dan *back-and-forth communication system* pada mesin *chatbot* menggunakan *Rapid Automatic Keywords Extraction*.
- c. Mengetahui akurasi jawaban yang dihasilkan oleh aplikasi *chatbot* untuk menjawab pertanyaan *user*.

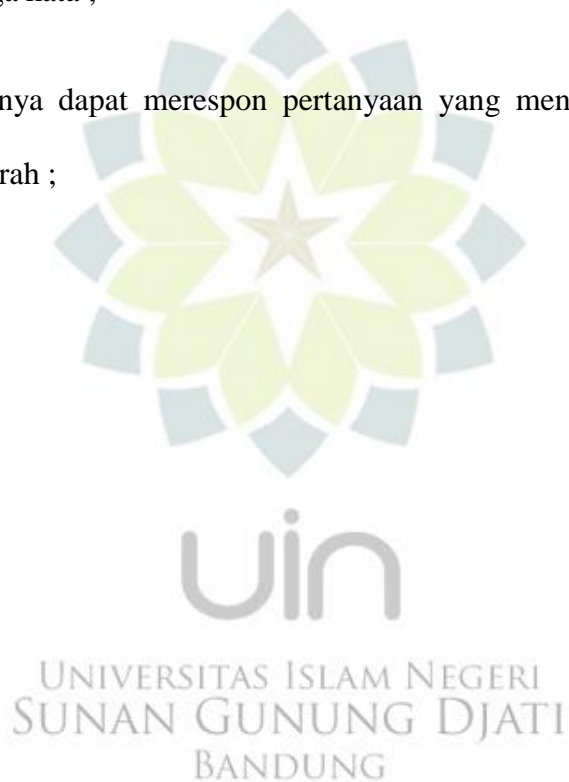
#### 1.4 Batasan Masalah

Karena adanya keterbatasan pada kemampuan yang dimiliki peneliti, baik segi waktu, biaya maupun pengetahuan, maka penelitian ini dibatasi. Berikut merupakan batasan masalah dari aplikasi yang dibuat, yaitu :

- a. Data yang digunakan merupakan hukum *taharah* menurut madzhab Imam Syafi'i yang bersumber dari buku 125 Masalah Taharah karya Anis Sumaji, Fikih Taharah karya Dr. Yusuf Al-Qardhawi dan Ilmu Fiqih (Safinatunnaja) Berikut Penjelasannya karya Syekh Salim Ibnu Samir Al-Hadhrami ;
- b. Informasi yang dihasilkan oleh *chatbot* bersumber dari data yang telah diklasifikasi sebelumnya dan disimpan dalam *operational database* dengan format CSV dan JSON ;
- c. Distribusi Algoritma *Naive Bayes* yang digunakan untuk mengklasifikasi pertanyaan adalah distribusi *Multinomial Naive Bayes* ;
- d. Aplikasi *chatbot* hukum *taharah* merupakan aplikasi berbasis web ;
- e. Bahasa pemrograman yang digunakan adalah Python dengan Framework Flask ;
- f. Model yang diterapkan pada Aplikasi merupakan *Retirival-Based model*, bukan *Generative model* ;
- g. Aplikasi hanya melakukan percakapan menggunakan bahasa Indonesia ;

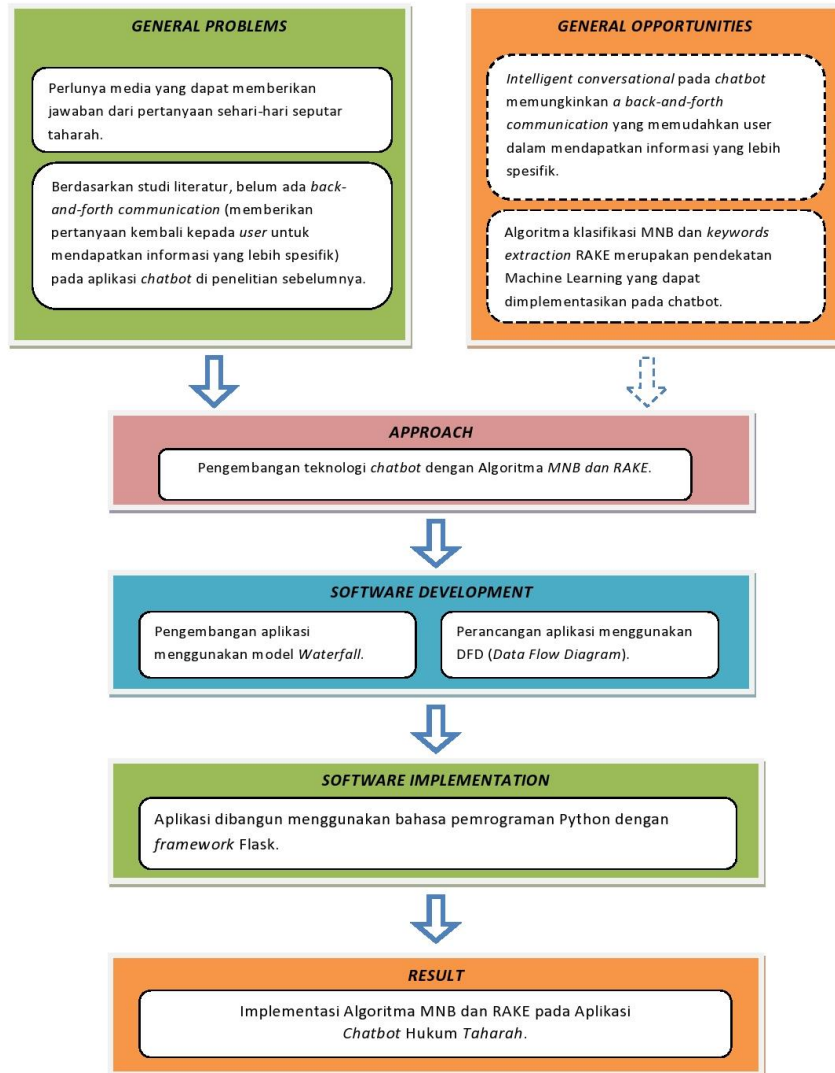


- h. Aplikasi hanya bisa merespon percakapan dalam bahasa alami dan tidak dapat melakukan perhitungan matematis ;
- i. Aplikasi tidak dapat membedakan lawan bicara berdasarkan identitasnya seperti jenis kelamin, umur atau nama ;
- j. Aplikasi hanya dapat merespon pertanyaan dengan jumlah minimum sebanyak tiga kata ;
- k. Aplikasi hanya dapat merespon pertanyaan yang mengandung *keyword* tentang taharah ;



## 1.5 Kerangka Pemikiran

Gambar 1.5 merupakan kerangka pemikiran dari penelitian tugas akhir ini.



**Gambar 1.5** Kerangka Pemikiran

Gambar diagram diatas menggambarkan garis besar alur logika penelitian yang dilakukan. Teknologi yang dikembangkan pada penelitian ini berupa aplikasi chatbot dengan *Multinomial Naïve Bayes* dan *Rapid Automatic Keywords Extraction* sebagai *proposed method* atau metode yang diusulkan. *Software*

*development model* yang digunakan adalah model *Waterfall*, sedangkan *software design* yang digunakan adalah *Data Flow Diagram* atau DFD. Kemudian *chatbot* dibangun menggunakan bahasa pemrograman Python dengan *framework* Flask.

## 1.6 Metodologi Penelitian

### 1.6.1 Pengumpulan Data

Pengumpulan data pada penelitian ini dilakukan menggunakan teknik Optical Character Recognition. Optical Character Recognition (OCR) adalah proses konversi gambar huruf menjadi karakter ASCII yang dikenali oleh komputer[9]. Gambar huruf yang dimaksud dapat berupa hasil *scan* dokumen, hasil *print-screen* halaman web, hasil foto dan lain-lain [9]. Teknik OCR digunakan untuk *scan* buku sebagai bahan data training. Buku yang digunakan adalah buku-buku fiqih Islam, diantaranya sebagai berikut :

- 1.) 125 Masalah Taharah. Buku ini ditulis oleh Anis Sumaji dan diterbitkan pada tahun 2008 oleh PT Tiga Serangkai. Beliau penulis yang sekaligus penceramah mencoba membahas permasalahan yang sering dihadapi para jemaahnya di buku ini [10]. Mayoritas data diambil dari buku ini.
- 2.) Fiqih Taharah. Buku ini ditulis oleh Dr. Yusuf Al-Qardhawi yang diterjemahkan oleh Samson Rahman, MA dan diterbitkan oleh Pustaka Al-Kautsar pada tahun 2006. Buku ini menjadi istimewa karena penulis adalah ulama yang sudah diakui kepakarannya di dunia Islam baik masalah fikih, akidah dan lainnya[11].
- 3.) Ilmu Fiqih (Safinatunnaja) Berikut Penjelasannya. Buku ini ditulis oleh Syekh Salim Ibnu Samir Al-Hadhrami yang diterjemahkan oleh KH.Moch Anwar dan

diterbitkan pada tahun 2011 oleh Sinar Baru Algesindo. Penulis sengaja menerjemahkan kitab Safinatunnaja ini sebab, kitab yang sangat banyak dikaji dimana-mana, baik di pesantren-pesantren, majelis ta'lim maupun surau-surau[12].

#### 1.6.2 Perancangan dan Pemodelan

Tahap perancangan dilakukan setelah tahap analisis selesai. Pada tahapan ini dilakukan beberapa perubahan kebutuhan-kebutuhan kedalam desain mekanik, diantaranya DFD (*Data Flow Diagram*), *Flowchart* dan *Mock Up*.

#### 1.6.3 Penulisan Program

Pada tahapan ini dilakukan penulisan program sesuai dengan perancangan yang sudah dilakukan. Penulisan program menggunakan bahasa pemrograman *Python*.

#### 1.6.4 Pengujian Sistem

Pada tahapan ini, dilakukan pengujian pada aplikasi yang sudah dibuat menggunakan *black box testing*.

### 1.7 Sistematika Penulisan

Penulisan tugas akhir ini dibagi ke dalam enam bab, dengan susunan sistematika penulisan berikut:

#### **BAB I PENDAHULUAN**

Bab ini menjelaskan latar belakang masalah, perumusan masalah, batasan-batasan masalah, tujuan dan manfaat penelitian, metodologi penelitian dan sistematika penulisan.

## **BAB II LANDASAN TEORI**

Bab ini berisi teori-teori beserta penjelasannya yang akan digunakan sebagai tinjauan untuk memecahkan masalah dan mendukung penelitian.

## **BAB III ANALISIS DAN PERANCANGAN**

Bab ini menjelaskan hasil analisis masalah, data dan algoritma yang menjadi landasan untuk perancangan sistem. Bab ini juga membahas arsitektur sistem yang akan dibuat.

## **BAB IV IMPLEMENTASI DAN PENGUJIAN**

Bab ini menerangkan hasil implementasi dan hasil pengujian sistem yang telah dibangun.

## **BAB V PENUTUP**

Bab ini berisi pernyataan singkat yang menjelaskan kesimpulan dari penelitian yang dilakukan secara keseluruhan. Bab ini juga berisi saran untuk pengembangan penelitian yang lebih baik lagi kedepannya.

## **DAFTAR PUSTAKA**

Daftar Pustaka berisi sumber-sumber baik cetak maupun tertulis yang digunakan dalam penelitian dan dikutip dalam penyusunan.

## **LAMPIRAN**

Pada lampiran ini berisi dokumen-dokumen yang telah digunakan dalam proses penyusunan dan juga perancangan seperti *source code* serta kelengkapan dokumen lainnya.