IMPLEMENTASI METODE TF-IDF PADA APLIKASI *CHATBOT* BERBASIS *WEB* DALAM SISTEM PENJUALAN *ONLINE* (STUDI KASUS : PERCETAKAN RADJA PRINTING)

PROPOSAL SKRIPSI

Oleh:

JOSHUA DAVIAN KRISTANTO NIM. 1641720027



PROGRAM STUDI TEKNIK INFORMATIKA

JURUSAN TEKNOLOGI INFORMASI

POLITEKNIK NEGERI MALANG

2020

IMPLEMENTASI METODE TF-IDF PADA APLIKASI *CHATBOT* BERBASIS *WEB* DALAM SISTEM PENJUALAN *ONLINE* (STUDI KASUS : PERCETAKAN RADJA PRINTING)

**PROPOSAL SKRIPSI**

Digunakan Sebagai Syarat Maju Ujian Diploma IV

Politeknik Negeri Malang

Oleh:

JOSHUA DAVIAN KRISTANTO NIM. 1641720027



PROGRAM STUDI TEKNIK INFORMATIKA

JURUSAN TEKNOLOGI INFORMASI

POLITEKNIK NEGERI MALANG

2020

HALAMAN PENGESAHAN

IMPLEMENTASI METODE TF-IDF PADA APLIKASI *CHATBOT* BERBASIS *WEB* DALAM SISTEM PENJUALAN *ONLINE* (STUDI KASUS : PERCETAKAN RADJA PRINTING)

Disusun oleh:

JOSHUA DAVIAN KRISTANTO NIM. 1641720027

**Proposal Skripsi ini telah diuji pada 15 Januari 2020**

**Disetujui oleh:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 1. | Penguji I | : | Kolonel Lek. Dr. Arwin Datumaya Wahyudi Sumari  NRP. 515561 | ........................... |
|  |  |  |  |  |
| 2. | Penguji II | : | Deasy Sandhya Elya Ikawati,  S. Si., M. Si.  NIDN. 8129002 | ........................... |
|  |  |  |  |  |
| 4. | Pembimbing | : | Dimas Wahyu Wibowo, ST., MT  NIP. 19841009 201504 1 001 | ........................... |

Mengetahui,

|  |  |
| --- | --- |
| Ketua Jurusan  Teknologi Informasi | Ketua Program Studi  Teknik Informatika |
| Rudy Ariyanto, S.T., M.CS  NIP. 19711110 199903 1 002 | Imam Fahrur Rozi, ST., MT.  NIP. 198406102008121004 |
|  |  |

**DAFTAR ISI**

HALAMAN JUDUL ii

HALAMAN PENGESAHAN iii

DAFTAR ISI iv

[1. Judul Skripsi 1](#_Toc30673683)

[2. Latar Belakang 1](#_Toc30673684)

[3. Rumusan Masalah 2](#_Toc30673685)

[4. Batasan Masalah 2](#_Toc30673686)

[5. Tujuan 3](#_Toc30673687)

[6. Landasan Teori 3](#_Toc30673688)

[6.1 Penelitian Terdahulu 3](#_Toc30673689)

[6.2 *Chatbot* 6](#_Toc30673690)

[6.3 CodeIgniter 6](#_Toc30673691)

[6.4 XAMPP 6](#_Toc30673692)

[6.5 *MySQL* 7](#_Toc30673693)

[6.6 *Hypertext Pre-Processor* (PHP) 7](#_Toc30673694)

[6.7 *JavaScript* 7](#_Toc30673695)

[6.8 *Natural Language Processing* 8](#_Toc30673696)

[6.9 *Term Frequency-Inverse Document Frequency* (TF-IDF) 8](#_Toc30673697)

[6.10 *Artificial Intelligence Markup Language* (AIML) 9](#_Toc30673698)

[6.11 *Recall* dan *Precission* 9](#_Toc30673699)

[7. Metodologi Penelitian 10](#_Toc30673700)

[7.1 Data 10](#_Toc30673701)

[7.2 Metode Pengambilan Data 11](#_Toc30673702)

[7.3 Metode Pengolahan Data 11](#_Toc30673703)

[7.4 Metode Pengembangan Perangkat Lunak 12](#_Toc30673704)

[7.4.1 Analisis Kebutuhan Sistem 13](#_Toc30673705)

[7.4.2 Desain Sistem dan Aplikasi 13](#_Toc30673706)

[7.4.3 Implementasi Sistem 17](#_Toc30673707)

[7.4.4 Pengujian Sistem 17](#_Toc30673708)

[7.5 Metode Pengujian 17](#_Toc30673709)

[8. Jadwal Kegiatan 19](#_Toc30673710)

[DAFTAR PUSTAKA 21](#_Toc30673711)

# Judul Skripsi

Implementasi Metode TF-IDF Pada Aplikasi Chatbot Berbasis Web Dalam Sistem Penjualan Online Radja Printing.

# Latar Belakang

Tidak bisa dipungkiri, teknologi mengalami perkembangan yang pesat seiring dengan berjalannya waktu. Perkembangan ini tentunya mempengaruhi banyak bidang, salah satunya dalam bidang perekonomian. Hal ini membuat para pelaku binis memanfaatkan teknologi yang ada untuk meningkatkan pemasaran produk mereka. Dari yang semula aktivitas jual-beli dilakukan secara manual, mulai digantikan dengan teknologi yang canggih. Hal ini memudahkan calon pembeli untuk melakukan transaksi melalui akses *website*.

Berdasarkan penelitian sebelumnya terkait penelitian ini adalah penelitian milik Dhebys Suryani Hormansyah dan Yoga Putera Utama yang dilakukan pada tahun 2018, penelitian tersebut menggunakan *chatbot* untuk layanan publik kesehatan di Kota Malang dengan metode *Term Frequency-Inverse Document Frequency* (TF-IDF). Sistem *chatbot* pada aplikasi tersebut menggantikan peran *customer service* dikarenakan pada Sistem Informasi Layanan Publik Kesehatan tidak terdapat layanan atau fitur pencarian, sehingga pengguna aplikasi kesulitan dalam pencarian data. Pada penelitian tersebut menggunakan metode TF-IDF, dimana metode ini digunakan untuk menghitung kata kunci dalam setiap data yang tersedia dalam Sistem Informasi Layanan Publik Kesehatan Malang.

Gerai Percetakan Radja Printing merupakan usaha percetakan yang berdiri tahun 2016. Berlokasi di kota Probolinggo, gerai percetakan ini bergerak di bidang percetakan dan penjualan *banner*. Permasalahan yang dihadapi oleh usaha ini yaitu tidak adanya sistem yang memungkinkan pembeli untuk melakukan pemesanan secara *online*. Permasalahan lain yang dihadapi oleh usaha ini yaitu adanya keterbatasan jam kerja *customer service* dalam merespon pertanyaan yang diajukan oleh calon pembeli. Dikarenakan usaha ini memiliki jam kerja yang dibatasi oleh aturan perusahaan, sedangkan *customer* seringkali membutuhkan informasi yang cepat dan tepat.

TF-IDF merupakan metode algoritma yang menentukan frekuensi dari kemunculan sebuah term dalam dokumen yang bersangkutan. Metode ini untuk menghitung nilai TF dan IDF pada setiap kata di setiap dokumen. Semakin besar jumlah kemunculan suatu term (TF tinggi) dalam dokumen, semakin besar pula bobotnya atau akan memberikan nilai kesesuaian yang semakin besar. Berdasarkan pada studi kasus yang ada, penelitian ini menerapkan metode TF-IDF dalam Sistem Informasi Penjualan Online Gerai Radja Printing, yang diharapkan memudahkan serta menyelesaikan permasalahan yang telah dipaparkan diatas. Aplikasi *chatbot* ini diharapkan dapat menjawab pertanyaan *customer* dengan penjelasan yang mudah dipahami.

# Rumusan Masalah

Berdasarkan latar belakang di atas, maka rumusan masalah yang dapat diambil adalah sebagai berikut:

1. Merancang dan membuat sistem aplikasi *chatbot* yangdapat digunakan untuk menggantikan peran *customer service* Radja Printing.
2. Mengukur kinerja aplikasi *chatbot* menggunakan *Recall* dan *Precision* sebagai tolak ukur dari ketepatan respon *chatbot* yang dirancang.

# Batasan Masalah

Agar skripsi yang berjudul Implementasi Metode TF-IDF Pada Aplikasi *Chatbot* Berbasis Web Pada Sistem Penjualan Online Radja Printing dapat berjalan sesuai rencana dan tujuan awal, maka penulis menetapkan batasan-batasan masalah yaitu:

Informasi mengenai layanan yang tersedia.

Informasi mengenai jam layanan percetakan.

Informasi mengenai harga layanan percetakan.

# Tujuan

Tujuan dari pembuatan sistem ini adalah sebagai berikut :

Untuk memudahkan pelanggan saat ingin bertanya mengenai informasi percetakan dengan jawaban sesuai jam bertanya dan pertanyaan beruntut akan menghasilkan jawaban yang beruntut.

Efisiensi dari proses bisnis percetakan yang dimana membantu kinerja dari *customer service* percetakan untuk melayani pertanyaan dari *customer* selama 24 jam.

# Landasan Teori

Tinjauan pustaka merupakan bagian yang akan membahas tentang penyelesaian masalah yang akan memberikan jalan keluarnya. Dalam hal ini akan dikemukakan beberapa teori-teori yang berkaitan dengan masalah yang diangkat.

## Penelitian Terdahulu

Penelitian terdahulu ini menjadi salah satu acuan penulis dalam melakukan penelitian sehingga penulis dapat memperkaya teori yang digunakan dalam mengkaji penelitian yang dilakukan. Dari penelitian terdahulu, penulis tidak menemukan penelitian dengan judul yang sama dengan judul penelitian penulis. Namun, penulis mengangkat beberapa penelitian sebagai referensi dalam memperkaya kajian pada penelitian penulis. Berikut merupakan penelitian terdahulu berupa beberapa jurnal terkait dengan penelitian yang akan dilakukan.

Tabel 6.1.1 Penelitian Terdahulu

|  |  |  |
| --- | --- | --- |
| **Nama Peneliti** | **Judul Penelitian** | **Hasil Penelitian** |
| Suryani, Dhebys & Larasati Amalia, Eka, 2017. | Aplikasi Chatbot Objek Wisata Jawa Timur Berbasis AIML. | Aplikasi *chatbot* dapat memberikan informasi kepada wisatawan yang ingin berwisata di wilayah Jawa Timur |
| Suryani, Dhebys & Putera, Yoga, 2018. | Aplikasi Chatbot Berbasis Web Pada Sistem Informasi Layanan Publik Kesehatan di Malang dengan Menggunakan Metode TF-IDF. | Aplikasi *chatbot* menggunakan metode TF-IDF dan *Cosine Similarity* untuk mencari jawaban pada sistem. |
| Suryani, Dhebys & Aulia, Indinabilah, 2018. | Penerapan Metode TF-IDF dan N-Gram pada Pengembangan Aplikasi Chatbot Berbasis LINE untuk Layanan Publik Kesehatan di Kota Malang. | *Question-Answering* dalam bentuk *chatbot* menggunakan N-Gram, TF-IDF dan *Cosine Similarity* dapat berkomunikasi dan menyampaikan informasi. |
| Tirtana, A., Zulkarnain, A., Dwi Listio, Y., 2019 | Pembuatan Sistem Pencarian Pekerjaan Menggunakan TF-IDF. | Penerapan metode TF-IDF memberikan hasil pencarian yang lebih relevan daripada pencarian tanpa pembobotan. |
| Hiroshi, Yamaguchi & Mozgoyov, Maxim, 2018 | A Chatbot Based on AIML Rules Extracted From Twitter Dialogues. | Kombinasi antara TF-IDF dan AIML memberikan hasil respon *chatbot* yang sesuai. |
| Astiningrum, Mungki & Shoburu Rohmah, Maya, 2018 | Implementasi NLP dengan Konversi Kata pada Sistem Chatbot Konsultasi Laktasi. | Diterapkannya metode *Levenshtein Distance* dan TF-IDF dan *Cosine Similarity* membuat aplikasi *chatbot* layak digunakan untuk *customer service* pusat laktasi. |
| Natadian Astuti, Rani & Fatchan, Muhammad, 2019 | Perancangan Aplikasi Chatbot untuk Industri Komersial 4.0 | Adanya aplikasi *chatbot* membuat peran dari *customer service* menjadi lebih efektif karena dapat melayani pertanyaan dari *customer* selama 24 jam. |
| Satria Paliwahet, I Nyoman & Gede Darma Putra, I Ketut, 2017 | Pencarian Informasi Wisata Daerah Bali Menggunakan Teknologi Chatbot. | Sistem *chatbot* dengan menggunakan metode *Fulltext Search Boolean Mode* dari MySQL dapat diterapkan dengan baik. |
| Melita, Ria & Dirjam, Taslimun, 2018 | Penerapan Metode Term Frequency – Inverse Document Frequency (TF-IDF) dan Cosine Similarity pada Sistem Temu Kembali Informasi untuk Mengetahui Syarah Hadits Berbasis Web (Studi Kasus : Syarah Umdatil Ahkam) | Metode TF-IDF dan *Cosine Similarity* berhasil diterapkan dengan memberikan hasil berupa *output* dokumen, yaitu *syarah hadits* sesuai dengan *query* yang di-input-kan. |
| Riyani, Ade & Burhanuddin Aulia, 2019 | Penerapan Cosine Similarity dan Pembobotan TF-IDF untuk Mendeteksi Kemiripan Dokumen. | Algoritma *Cosine Similarity* dan pembobotan TF-IDF telah berhasil mendeteksi kemiripan pada suatu dokumen. |

Dari beberapa judul penelitian yang telah dipaparkan, terpilih penelitian yang dilakukan oleh Dhebys Suryani dan Yoga Putera yang berjudul “Aplikasi Chatbot Berbasis Web Pada Sistem Informasi Layanan Publik Kesehatan di Malang dengan Menggunakan Metode TF-IDF” sebagai referensi utama penulis dalam melakukan penelitian. Hal ini dikarenakan adanya persamaan metode yang digunakan oleh penulis, yaitu TF-IDF dan *Cosine Similarity*, sebagai metode utama dalam pembuatan aplikasi *chatbot*.

## *Chatbot*

*Chatbot* adalah program komputer yang dapat melakukan percakapan melalui media tulisan. Percakapan dapat terjadi dengan manusia atau chatbot yang lain. Secara harfiah *chatbot* berasal dari dua kata, yaitu *chat* dan *bot*. Dalam dunia komputer, chat dapat diartikan sebagai kegiatan komunikasi yang menggunakan sarana tulisan. Sedangkan *bot* merupakan program yang memiliki sejumlah data yang bila diberikan input akan menghasilkan output sebagai jawaban.

## CodeIgniter

CodeIgniter merupakan sebuah *framework* yang dibuat dengan menggunakan bahasa *Hypertext Pre-Processor* (PHP), yang dapat digunakan untuk pengembangan web secara cepat. Adapun *framework* sendiri dapat diartikan sebagai suatu struktur pustaka-pustaka, kelas-kelas dan infrastruktur *run-time* yang dapat digunakan oleh programmer untuk mengembangkan aplikasi web secara cepat. Tujuan penggunaan *framework* adalah untuk mempermudah pengembang *web* mengembangkan aplikasi *web* yang *robust* secara cepat tanpa kehilangan fleksibilitas.

## XAMPP

XAMPP adalah perangkat lunak bebas, yang mendukung banyak sistem operasi, merupakan kompilasi dari beberapa program. Fungsinya adalah sebagai *server* yang berdiri sendiri (*localhost*), yang terdiri atas program *Apache HTTP Server*, MySQL *database*, dan penerjemah bahasa yang ditulis dengan bahasa pemrograman PHP dan Perl. Nama XAMPP merupakan singkatan dari X (empat sistem operasi apapun), Apache, MySQL, PHP dan Perl. Program ini tersedia dalam *General Public License* dan bebas, merupakan web server yang mudah digunakan yang dapat melayani tampilan halaman web yang dinamis

## *MySQL*

MySQL adalah salah satu jenis database server yang sangat terkenal. Kepopulerannya disebabkan MySQL menggunakan SQL sebagai bahasa dasar untuk mengakses databasenya. MySQL bersifat free dengan lisensi GNU General Public License (GPL). Dengan adanya keadaan ini maka anda dapat meggunakan software ini dengan bebas tanpa perlu harus takut dengan lisensi yang ada.MySQL termasuk jenis *Relational Database Management System* (RDBMS). Itulah sebabnya istilah tabel, baris, kolom digunakan pada MySQL. Pada MySQL sebuah database mengandung satu atau sejumlah table

## *Hypertext Pre-Processor* (PHP)

*PHP* singkatan dari  *Hypertext Preprocessor* yaitu bahasa pemrograman web server-side yang bersifat open source. PHP merupakan script yang terintegrasi dengan *Hyper Text Markup Languange* (HTML) dan berada pada *server* (*server side HTML embedded scripting*). PHP adalah script yang digunakan untuk membuat halaman website yang dinamis. Dinamis berarti halaman yang akan ditampilkan dibuat saat halaman itu diminta oleh *client*. Mekanisme ini menyebabkan informasi yang diterima client selalu terbaru/*up to date*. Semua *script* PHP dieksekusi pada *server* dimana *script* tersebut dijalankan.

## *JavaScript*

JavaScript adalah bahasa pemrograman web yang bersifat *Client Side Programming Language*. *Client Side Programming Language* adalah tipe bahasa pemrograman yang pemrosesannya dilakukan oleh *client*. Aplikasi *client* yang dimaksud merujuk kepada *web browser* seperti Google Chrome, Mozilla Firefox, Opera Mini dan sebagainya. JavaScript pertama kali dikembangkan pada pertengahan dekade 90’an. Meskipun memiliki nama yang hampir serupa, JavaScript berbeda dengan bahasa pemrograman Java. Untuk penulisannya, JavaScript dapat disisipkan di dalam dokumen HTML ataupun dijadikan dokumen tersendiri yang kemudian diasosiasikan dengan dokumen lain yang dituju. JavaScript mengimplementasikan fitur yang dirancang untuk mengendalikan bagaimana sebuah halaman web berinteraksi dengan penggunanya (Henderson, 2009:256).

## *Natural Language Processing*

*Natural Language Processing* (NLP) adalah bidang penelitian dan aplikasi yang mengeksplorasi bagaimana komputer dapat digunakan untuk memahami dan memanipulasi teks bahasa alami. Peneliti NLP bertujuan untuk mengumpulkan pengetahuan tentang bagaimana manusia memahami dan menggunakan bahasa sehingga peralatan dan teknik pemasangan dapat dikembangkan untuk membuat sistem komputer memahami dan memanipulasi bahasa alami untuk melakukan tugas yang disukai, (Indana, et al 2018).

## *Term Frequency-Inverse Document Frequency* (TF-IDF)

Metode TF -IDF *adalah* cara pemberian bobot hubungan suatu kata (term) terhadap dokumen.

1. TF murni (raw TF), nilai TF diberikan berdasarkan jumlah kemunculan suatu term di dokumen.

(6.1)

Ket: = *term frequency*

= Banyaknya kemunculan term pada dokumen

1. Inverse Document Frequency (IDF) merupakan sebuah perhitungan dari bagaimana term di distribusikan secara luas pada koleksi dokumen yang bersangkutan.IDF menunjukkan hubungan ketersediaan sebuah term dalam seluruh dokumen. Semakin sedikit jumlah dokumen yang mengandung term mengandung term yang dimaksud., maka nilai IDF semakin besar.

(6.2)

Ket: = *Inverse Document Frequency*

= Jumlah dokumen yang terambil dalam sistem

= Banyak dokumen dalam koleksi term *t* muncul di dalamnya

1. Jenis formula **TF** yang biasa digunakan untuk perhitungan adalah **TF murni (*raw* TF)**. Dengan demikian rumus umum untuk ***Term Weighting* TF -IDF** adalah penggabungan dari formula perhitungan ***raw* TF** dengan formula **IDF** dengan cara mengalikan nilai **TF** dengan nilai **IDF**:

(6.3)

Ket: = Bobot *term* terhadap dokumen

= Jumlah kemunculan *term*  dalam dokumen

= Jumlah semua dokumen yang ada di dalam *database*

= Jumlah dokumen yang mengandung *term*

## *Artificial Intelligence Markup Language* (AIML)

Artificial Intelligence Markup Language (AIML) adalah sebuah bahasa yang mendeskripsikan objek data dan perilaku program komputer yang memprosesnya. Objek AIML tersusun atas unit-unit yang disebut topics dan categories, berisi data yang sudah ter-parsing berisi karakter-karakter, beberapa diantaranya berupa data karakter, yang lainnya dapat berupa elemen AIML. Elemen AIML mengkapsulasi pengetahuan dalam bentuk stimulus-response di dokumen.

## *Recall* dan *Precission*

Menurut Purwono (2010) Recall (perolehan berhubungan dengan kemampuan sistem untuk memanggil dokumen yang relevan. sedangkan precision (ketepatan) adalah berkaitan dengan kemampuan sistem untuk tidak memanggil dokumen yang tidak relevan. *Recall* adalah proporsi jumlah dokumen yang dapat ditemukan-kembali oleh sebuah proses pencarian di sistem IR. Rumusnya: Jumlah dokumen relevan yang ditemukan dibagi dengan jumlah semua dokumen relevan di dalam koleksi. *Precision* adalah proporsi jumlah dokumen yang ditemukan dan dianggap relevan untuk kebutuhan si pencari informasi. Rumusnya: Jumlah dokumen relevan yang ditemukan dibagi dengan jumlah semua dokumen yang ditemukan. Sedangkan *Precision* dapat diartikan sebagai kepersisan atau kecocokan (antara permintaan informasi dengan jawaban terhadap permintaan itu). Jika seseorang mencari informasi di sebuah sistem, dan sistem menawarkan beberapa dokumen, maka kepersisan ini sebenarnya juga adalah relevansi. Artinya, seberapa persis atau cocok dokumen tersebut untuk keperluan pencari informasi, bergantung pada seberapa relevan dokumen tersebut bagi si pencari. Rumus penilaian*precision* yang digunakan dalam penelitian ini dituliskan dalam rumus 6.4

(6.4)

Nilai relatif *Recall* dihitung dengan rumus 6.5.

(6.5)

# Metodologi Penelitian

## Data

Data-data yang dibutuhkan dalam pembuatan aplikasi ini meliputi, data produk dari gerai percetakan Radja Printing dan data percakapan antara *customer* *service* dengan *customer* yang akan menjadi contoh dari *chatbot.* Data barang yang ada dalam gerai percetakan Radja Printing disediakan dalam tabel 7.1.1.

Tabel 7.1.1 Data Barang

|  |  |  |  |
| --- | --- | --- | --- |
| **No.** | **Nama** | **Kategori** | **Harga** |
| 1 | Banner Biasa | Banner | 18000 |
| 2 | Banner Glosy | Banner | 25000 |
| 3 | Banner Korea | Banner | 30000 |
| 4 | Stiker Cina | Stiker | 55000 |
| 5 | Stiker Ritrama | Stiker | 75000 |
| 6 | Stiker Luster | Stiker | 60000 |
| 7 | Stiker Cloth | Stiker | 45000 |
| 8 | Stiker Oneway | Stiker | 75000 |
| 9 | Stiker Transparan | Stiker | 80000 |
| 10 | Art Paper 120 | Art Paper | 3500 |
| 11 | Art Paper 150 | Art Paper | 3500 |
| 12 | Art Paper 230 | Art Paper | 3500 |
| 13 | Art Paper 260 | Art Paper | 4000 |
| 14 | Art Paper Bontax | Art Paper | 4500 |
| 15 | Art Paper Jasmine | Art Paper | 4500 |
| 16 | Art Paper BC TIK | Art Paper | 5000 |
| 17 | Art Paper Vinil Susu | Art Paper | 8000 |
| 18 | Art Paper Vinil Transparan | Art Paper | 8000 |
| 19 | X Banner | Banner | 40000 |
| 20 | Roll Banner | Banner | 180000 |
| 21 | Pigura A3 | Pigura | 45000 |
| 22 | Pigura A4 | Pigura | 60000 |
| 23 | Brosur | Brosur | 175000 |

Sedangkan untuk data pertanyaan yang biasa ditanyakan oleh *customer* disediakan dalam tabel 7.1.2.

Tabel 7.1.2 Data Pertanyaan

|  |
| --- |
| **Pertanyaan** |
| Apa saja produk yang ada di percetakan? |
| Minta daftar harga cetak banner |
| Minta daftar harga produk di toko ini |
| Mau cetak banner |
| Selamat Pagi |
| Halo |
| Dimana alamat toko? |
| Jam berapa toko buka? |
| Berapa lama cetak banner korea |
| Siapa nama pemilik toko? |

## Metode Pengambilan Data

Metode yang digunakan dalam pengambilan data adalah melalui wawancara*.* Wawancara dilakukan menggunakan media *Whatsapp* dengan pemilik gerai percetakan. Wawancara dilakukan pada tanggal 20 November 2019 di tempat percetakan itu sendiri yang berlokasi di kota Probolinggo. Hasil wawancara yang dilakukan adalah didapatkan data barang yang ada di percetakan, contoh-contoh pertanyaan yang biasa diajukan *customer*, dan informasi-informasi lain terkait dengan operasional percetakan.

## Metode Pengolahan Data

1. Tahap Pertama

Pada tahap pertama ditentukan banyak kalimat yang akan digunakan sebagai *document.* Kalimat yang akan digunakan menggunakan *Natural Language Processing* (NLP) dimana NLP membantu komputer untuk mengolah kalimat yang digunakan manusia untuk berkomunikasi satu sama lain. *Term* merupakan kata kunci dari sebuah kalimat.

1. Tahap Kedua

Pada tahap kedua dilakukan pengumpulan hasil dari jumlah frekuensi *term* pada *document*. Penghitungan *IDF* yaitu log jumlah document dibagi dengan jumlah *document* yang mengandung *term*. Setelah mengetahui *raw TF* dan *IDF* berikutnya kedua nilai dikalikan dan akan menghasilkan *Term Weighting TF-IDF.* TF-IDF mengandung *term* yang memilki bobot. Penghitungan bobot dengan menghitung jumlah *term* yang ada pada *document.*

1. Tahap Ketiga

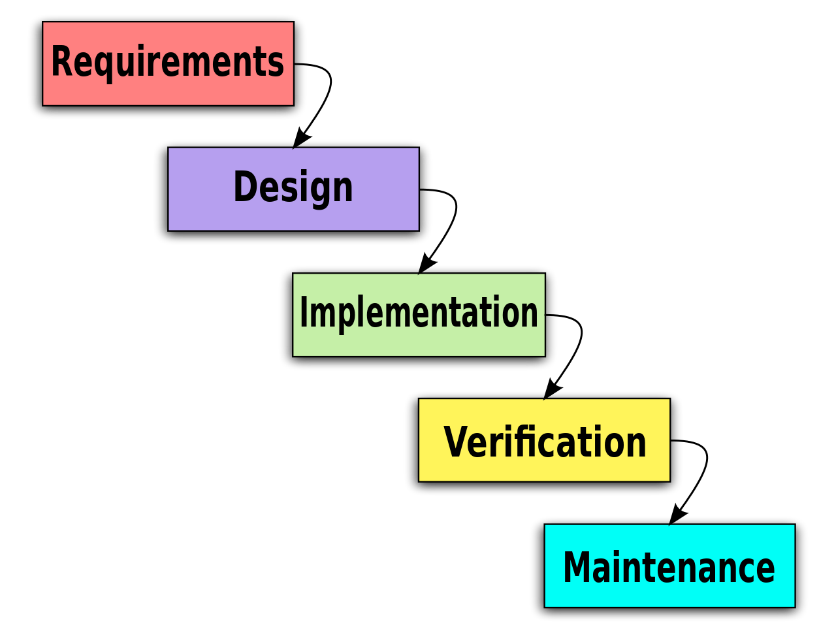
Pada tahap ketiga dimulai penggunaan AIML sebagai perhitungan kemiripan antar kalimat serta mengukur kemiripan teks yang popular. Kemiripan antar kalimat ini yang menjadi acuan untuk menentukan apa keinginan pengguna dan sistem akan mengambil kata kunci sebagai acuan untuk menampilkan jawaban sesuai kalimat yang ditulis. Kemudian akan dilakukan proses Recall dan Precission untuk menghitung keakuratan jawaban dari pertanyaan yang diberikan.

1. Tahap Keempat

Pada tahap keempat dilakukan implementasi, *testing* dan *maintenance* dari hasil kemiripan kata dengan kata kunci yang muncul untuk menentukan jawaban apa yang diinginkan dan sesuai dengan harapan dan rule berjalannya sistem. Untuk melakukan implementasi, penulis menggunakan PHP sebagai alat bantu dalam menerapkan rumus dari NLP, TF-IDF, AIML, serta Recall dan Procession.

## Metode Pengembangan Perangkat Lunak

Metode yang digunakan dalam perancangan “Implementasi Metode TF-IDF pada Aplikasi *Chatbot* Berbasis *Web* dalam Sistem Penjualan Online Radja Printing” adalah *System Development Life Cycle* (SDLC) dengan *waterfall* model. Waterfall model merupakan bentuk umum yang digunakan dalam perancangan sebuah sistem karena dalam setiap tahapan yang dilakukan harus diselesaikan sebelum menuju tahap selanjutnya sehingga tahapan dilakukan secara berurutan dan mendapatkan hasil yang maksimal. Tahapan-tahapan yang dilakukan yaitu :



Gambar 7.1 Waterfall Model

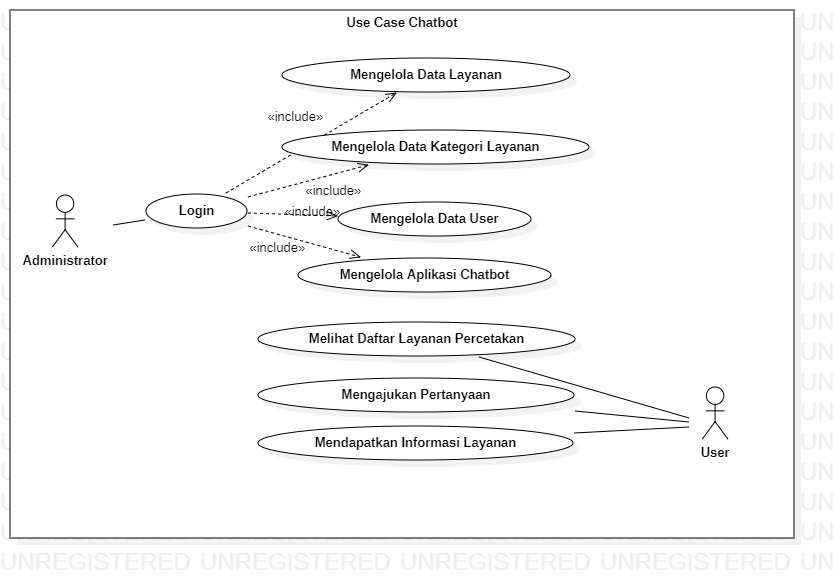
### Analisis Kebutuhan Sistem

Pada tahap ini dilakukan analisis kebutuhan sistem mencakup pengambilan data yang dibutuhkan sistem, kemudian nantinya sebagai pendukung/penunjang pengolahan data maupun penunjang pengerjaan. Adapun metode pengambilan data yang dibutuhkan sebagai pendukung dalam proses pembuatan aplikasi dengan studi literatur sebagai berikut :

1. Pengumpulan Data
2. Studi Pustaka
3. Observasi di Internet

### Desain Sistem dan Aplikasi

Pada tahap ini dilakukan penerjemahan kebutuhan sistem dalam sebuah perancangan sebelum dilakukan implementasi ke dalam bentuk *coding*. Tahap ini berfokus pada struktur data, arsitektur data, arsitektur perangkat lunak, dan representasi *interface.*



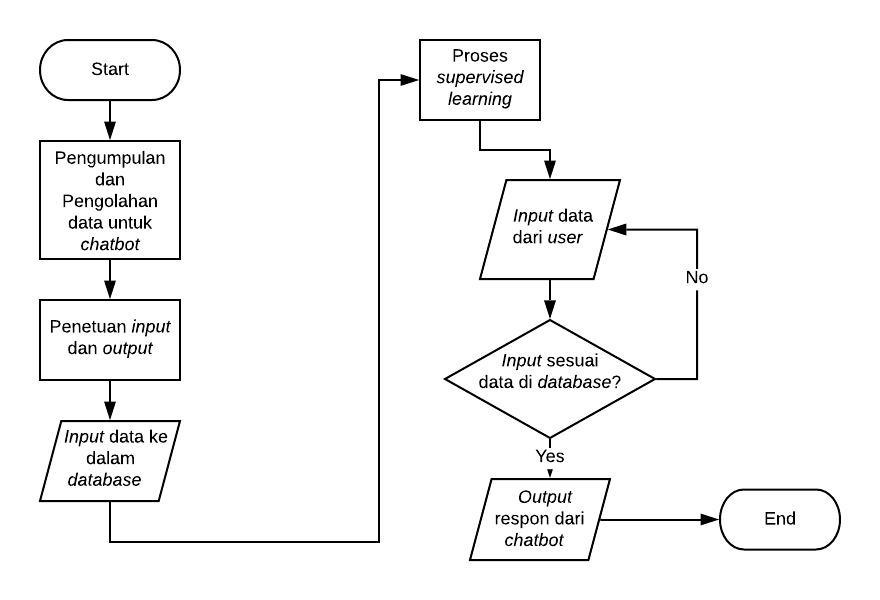
Gambar 7.2 Use Case Diagram

7.4.2.1 Penjelasan Use Case

Use case merupakan sebuah teknik yang digunakan dalam pengembangan sebuah *software* atau sistem informasi untuk menangkap kebutuhan fungsional dari sistem yang bersangkutan. Use case menjelaskan interaksi yang terjadi antara aktor, yaitu inisiator dari sistem itu sendiri, dengan sistem yang ada. Sebuah use case direpresentasikan dengan urutan langkah yang sederhana. Definisi dari aktor beserta deskripsinya dijelaskan dalam tabel 7.1.

Tabel 7.1 Definisi Aktor

|  |  |
| --- | --- |
| Aktor | Deskripsi |
| Admin | Admin adalah orang yang menggunakan aplikasi. Admin dapat melakukan aktifitas diantaranya mengelola data barang, mengelola data kategori, mengelola data transaksi, mengelola data user, mengelola data chatbot. Data chatbot terdiri dari CRUD pertanyaan dan CRUD jawaban. |
| User | User adalah pengguna jasa percetakan yang dapat mengajukan pertanyaan serta mendapat umpan balik dari sistem. |

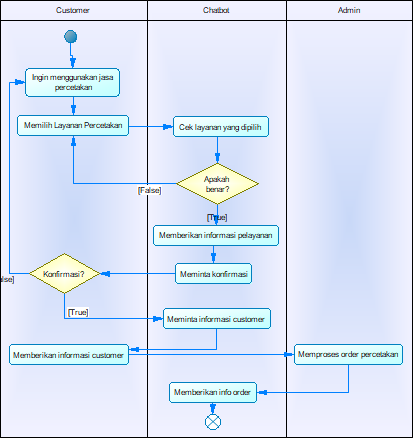


Gambar 7.3 Flowchart Sistem Aplikasi *Chatbot*

7.4.2.2 Penjelasan Flowchart Sistem Aplikasi *Chatbot*

Flowchart adalah suatu bagan dengan simbol-simbol tertentu yang menggambarkan urutan proses secara mendetail dan hubungan antara suatu proses (instruksi) dengan proses lainnya dalam suatu program.

Pada perancangan chatbot untuk studi kasus ini, *flowchacrt* dimulai dengan proses pengumpulan data dan pengolahan data, kemudian dilanjutkan dengan proses penentuan *input* dan *response* dari chatbot yang akan dibuat. Setelah itu dilanjutkan dengan proses *input* data ke dalam *database,* dan kemudian *chatbot* akan melakukan proses *supervised learning* untuk menambah pengetahuan pada sistem *chatbot* itu sendiri. *User* kemudian melakukan *input* pertanyaan pada sistem *chatbot*. Setelah itu, sistem akan memproses *input* pertanyaan dari *user,* apabila pertanyaan tersebut cocok dengan yang ada di database, maka sistem akan memberikan respon jawaban sesuai dengan *input* yang telah diberikan. Apabila tidak cocok, maka *user* akan diberikan pemberitahuan untuk melakukan *input* pertanyaan ulang.



Gambar 7.4 Activity Diagram

7.4.2.3 Penjelasan Activity Diagram

*Activity diagram* adalah bentuk visual dari alur kerja yang berisi aktivitas dan tindakan, yang juga dapat berisi pilihan, atau pengulangan. Dalam *Unified Modelling Language* (UML), diagram aktivitas dibuat untuk menjelaskan aktivitas komputer maupun aktivitas dari organisasi. Tujuan utama dari activity diagram adalah menggambarkan urutan aktivitas dalam sebuah proses agar lebih mudah dipahami. Pada gambar 7.4 diatas telah disediakan *activity diagram* dari sistem *chatbot* yang akan dibuat.

### Implementasi Sistem

Implementasi merupakan tahap pemrograman. Pada tahap ini program dibuat sesuai dengan kebutuhan dan fungsi yang dibutuhkan dan diinginkan. Namun selain itu juga program di analisis apakah telah sesuai dengan desain sistem yang dibuat. Proses implementasi dilakukan sampai seluruh fungsi yang telah dirancang dapat diimplementasikan seluruhnya ke dalam sistem *chatbot*. Sistem yang diimplementasikan berjalan pada *platform web*.

### Pengujian Sistem

Tahap pengujian program merupakan tahap yang dilakukan setelah implementasi pembuatan, dimana tahap ini bertujuan untuk mengetahui apakah sistem yang dibuat telah sesuai dengan sistem dan apakah ada kesalahan yang terjadi dari implementasi program. Pengujian dilakukan dengan metode *recall* dan *precision*, dimana metode ini digunakan untuk mengukur akurasi respon dari *chatbot*. Respon yang dikeluarkan *chatbot* berdasarkan *input* dari user dihitung menggunakan rumus 6.4 dan 6.5.

7.4.5Penerapan Sistem

Tahap *maintenance* merupakan tahap yang dilakukan setelah sistem yang dibuat sudah berjalan sebagaimana mestinya. Sehingga pada tahap ini merupakan tahapan pemeliharaan dan memperbaiki kesalahan apabila ditemukan suatu masalah yang tidak ditemukan pada tahap sebelumnya. Tujuan dari *maintenance* adalah agar sistem dapat terpelihara dengan baik sehingga kesalahan-kesalahan atau cacat pada sistem dapat terminimalisir.

## Metode Pengujian

Metode Pengujian chatbot dibutuhkan dalam menguji kinerja dari Chatbot yang dibangun. Pengujian ini dilakukan dengan cara menjalankan fitur yang tersedia dalam aplikasi dan melihat hasilnya sudah sesuai dengan yang diharapkan atau belum. Pengujian dengan metode TF-IDF dan AIML yang nantinya akan digunakan untuk sistem chatbot. Serta menggunakan validasi untuk menyesuaikan pertanyaan yang diberikan dan jawaban yang dilakukan dengan sistem.

# Jadwal Kegiatan

**Tabel 1. Jadwal Kegiatan**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| No | Kegiatan | November | | | | Desember | | | | | Januari | | | | Februari | | | | Maret | | | | April | | | | Mei | | | |
| 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 |
| **1** | **Usulan Judul** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | Pencarian Konsep Ide |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | Pengajuan Topik Skripsi |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | Pengajuan Dosen Pembimbing |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **2** | **PraProposal** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | Bimbingan Pembuatan Proposal |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | Seminar Proposal |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **3** | **Analisa** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | Analisa Sistem |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | Analisa Kebutuhan |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **4** | **Desain** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | Desain Package Diagram |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | Desain Sequence Diagram |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | Desain Mockup |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **5** | **Implementasi** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | Pembuatan Sistem Informasi |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | Pembuatan Chatbot pada Sistem |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **6** | **Pengujian** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | Pengujian Sistem |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **7** | **Laporan** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | Pembuatan Laporan |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

# DAFTAR PUSTAKA

Suryani, Dhebys & Larasati, Eka. (2017). Aplikasi Chatbot Objek Wisata Jawa Timur Berbasis AIML. *SMARTICS Journal: 3,* 47-54.

Suryani, Dhebys & Putera, Yoga. (2018). Aplikasi Chatbot Berbasis Web Pada Sistem Informasi Layanan Publik Kesehatan di Malang dengan Menggunakan Metode TF-IDF. *Jurnal Informatika Polinema, 4(3):* 224-228.

Suryani, Dhebys & Aulia, Indinabilah. (2018). Penerapan Metode TF-IDF dan N-Gram pada Pengembangan Aplikasi Chatbot Berbasis LINE untuk Layanan Publik Kesehatan di Kota Malang. *Jurnal Informatika Polinema, 5(1):* 7-11

Dwi Listio, Y., Zulkarnain, A., Tirtana, A. (2019). Pembuatan Sistem Pencarian Pekerjaan Menggunakan TF-IDF. *Jurnal Ilmiah Teknologi Informasi Asia, 13(2):* 91-100.

Yamaguchi, Hiroshi & Mozgoyov, Maxim. (2018). A Chatbot Based on AIML Rules Extracted From Twitter Dialogues. *Communication Paper Poznan, 17:* 37-42.

Astiningrum, Mungki & Shoburu Rohmah, Maya. (2018). Implementasi NLP dengan Konversi Kata pada Sistem Chatbot Konsultasi Laktasi. *Jurnal Informatika Polinema, 5(1):* 46-52

Natadian Astuti, Rani & Fatchan, Muhammad. (2019). Perancangan Aplikasi Teknologi Chatbot untuk Industri Komersial 4.0. *e-Prosiding SNasTekS 2019, 4:* 339-348

Satria Paliwahet, I Nyoman & Gede Darma Putra, I Ketut. (2017). Pencarian Informasi Wisata Daerah Bali Menggunakan Teknologi Chatbot. *Lontar Komputer, 8(3):* 144-153

Melita, Ria & Dirjam, Taslimun. (2018). Penerapan Metode Term Frequency Inverse Document Frequency dan Cosine Similarity pada Sistem Temu Kembali Informasi untuk Mengetahui Syarah Hadits Berbasis Web (Studi Kasus: Syarah Umdatil Ahkam). *Jurnal Teknik Informatika, 11(2):* 149-164

Riyani, Ade & Burhanuddin, Auliya. (2019). Penerapan Cosine Similarity dan Pembobotan TF-IDF untuk Mendeteksi Kemiripan Dokumen. *Jurnal Linguistik Komputasional, 2(1): 23-27*