

BAB II. LANDASAN TEORI

2.1 Sistem Informasi

Suatu sistem dapat didefinisikan sebagai suatu kesatuan yang terdiri dari dua atau lebih komponen atau subsistem yang berinteraksi untuk mencapai suatu tujuan. Sistem informasi yaitu suatu sistem yang menyediakan informasi untuk manajemen dalam mengambil keputusan dan juga untuk menjalankan operasional perusahaan, di mana sistem tersebut merupakan kombinasi dari orang-orang, teknologi informasi dan prosedur-prosedur yang terorganisasi. Biasanya suatu perusahaan atau badan usaha menyediakan semacam informasi yang berguna bagi manajemen.

Sistem informasi merupakan data-data yang telah dikumpulkan, dikelompokkan kemudian diolah menjadi sebuah informasi yang bermanfaat dan bernilai bagi penerimanya. Sistem informasi merupakan sistem yang berkemampuan dapat mengumpulkan serta mengelompokkan informasi dari berbagai sumber dengan menggunakan berbagai media sehingga dapat menampilkan informasi. (Yakub dan Herman, 2021).

2.2 Komisi Pemilihan Umum

Menurut Firmanzah (2010:55), “Komisi Pemilihan Umum (KPU) yaitu merupakan satu-satunya lembaga yang mempunyai kewenangan dalam menyelenggarakan Pemilu Legislatif, Pemilu Presiden, dan Pemilihan Kepala Daerah di Indonesia. Seluruh aspek yang berkaitan dengan pemilu menjadi tanggung jawab KPU dan bukan lembaga lainnya”. Dalam UU Nomor 15 tahun 2011 Pasal 1 ayat (6) dijelaskan bahwa Komisi Pemilihan Umum adalah lembaga penyelenggara pemilu yang bersifat nasional, tetap, dan mandiri yang bertugas melaksanakan pemilu. Dalam pasal ini juga dijelaskan mengenai KPU Provinsi dan KPU Kabupaten atau Kota. Ayat (7) pasal ini menjelaskan bahwa KPU Provinsi adalah penyelenggara pemilu yang bertugas melaksanakan pemilu di provinsi, sedang KPU Kabupaten/Kota adalah penyelenggara pemilu yang bertugas melaksanakan pemilu di kabupaten/kota (ayat (8)).

Komisi Pemilihan Umum adalah lembaga negara yang menyelenggarakan pemilihan umum di Indonesia, yakni meliputi Pemilihan Umum Anggota DPR/DPD/DPRD, Pemilihan Umum Presiden dan Wakil Presiden, serta Pemilihan Umum Kepala Daerah dan Wakil Kepala Daerah. Komisi Pemilihan Umum tidak dapat disejajarkan kedudukannya dengan lembaga-lembaga negara yang lain yang kewenangannya

ditentukan dan diberikan oleh UUD 1945. Bahkan nama Komisi Pemilihan Umum belum disebut secara pasti atau tidak ditentukan dalam UUD 1945, tetapi kewenangannya sebagai penyelenggara pemilihan umum sudah ditegaskan dalam Pasal 22E ayat (5) UUD 1945 yaitu Pemilihan umum diselenggarakan oleh suatu komisi pemilihan umum yang bersifat nasional, tetap, dan mandiri. Artinya, bahwa Komisi Pemilihan Umum itu adalah penyelenggara pemilu, dan sebagai penyelenggara bersifat nasional, tetap dan mandiri. (Zalukhu, 2021).

2.3 Pemilih Pemula

Pemilih pemula di Indonesia di bagi atas tiga kategori. Pertama, pemilih yang rasional, yakni pemilih yang benar-benar memilih partai berdasarkan penilaian dan analisis mendalam. Kedua, pemilih kritis emosional, yakni pemilih yang masih idealis dan tidak kenal kompromi. Ketiga, pemilih pemula, yakni pemilih yang baru pertama kali memilih karena usia mereka baru memasuki usia pemilih. Menurut pasal I ayat (22) UU No. 10 Tahun 2008, pemilih adalah warga negara Indonesia yang telah genap berumur 17 (Tujuh belas) tahun atau lebih atau sudah/belum kawin. Kemudian pasal 19 ayat (1 dan 2) UU No. 10 Tahun 2008 menerangkan bahwa pemilih yang mempunyai hak memilih adalah warga negara Indonesia yang didaftar oleh penyelenggara pilkada dalam daftar pemilih dan pada hari pemungutan suara telah genap berumur 17 (Tujuh belas) tahun atau lebih atau sudah pernah kawin.

Sistem Pengertian tersebut dapat di tarik simpulan bahwa pemilih pemula adalah warga negara yang di daftar oleh penyelenggara pilkada dalam daftar pilkada. Dan baru mengikuti pilkada (memberikan suara) pertama kali sejak pilkada yang diselenggarakan di Indonesia dengan rentang usia 17-21 tahun. Kelompok pemilih pemula ini biasanya mereka yang berstatus mahasiswa serta pekerja muda. Pemilih pemula dalam ritual demokrasi (pilkada) selama ini sebagai objek dalam kegiatan politik, yaitu mereka yang masih memerlukan pembinaan dan pengembangan ke arah pertumbuhan potensi dan kemampuannya ke tingkat yang lebih optimal agar dapat berperan dalam bidang politik. Kelompok pemilih pemula ini biasanya mereka yang berstatus mahasiswa serta pekerja muda. Pemilih pemula. dalam ritual demokrasi (pilkada) selama ini sebagai objek dalam kegiatan politik, yaitu mereka yang masih memerlukan pembinaan dan pengembangan ke arah pertumbuhan potensi dan kemampuannya ke tingkat yang lebih optimal agar dapat berperan dalam bidang politik. (H. Basuki Rachmat dan Esther, 2015)

Dari definisi di atas dapat di simpulkan bahwa ciri-ciri pemilih pemula yaitu:

1. Warga negara Indonesia dan pada hari pemungutan suara sudah berumur 17 tahun atau lebih atau sudah pernah kawin.
2. Baru mengikuti pilkada (memberikan suara) pertama kali sejak pemilu yang diselenggarakan di Indonesia dengan rentang usia 17-21 tahun.
3. Mempunyai hak memilih dalam penyelenggaraan pilkada 2024.

2.4 Website

Website adalah halaman informasi yang disediakan melalui jalur internet sehingga bisa diakses di seluruh dunia selama terkoneksi dengan jaringan internet. *Website* merupakan komponent atau kumpulan komponen yang terdiri dari teks, gambar, suara animasi sehingga lebih merupakan media informasi yang menarik untuk dikunjungi.

Secara terminologi, *website* adalah kumpulan dari halaman-halaman situs, yang biasanya terangkum dalam sebuah domain atau subdomain, yang tempatnya berada di dalam *World Wide Web* (WWW) di Internet. Sebuah halaman web adalah dokumen yang ditulis dalam format HTML (*Hyper Text Markup Language*), yang hampir selalu bisa diakses melalui HTTP, yaitu protokol yang menyampaikan informasi dari server *website* untuk ditampilkan kepada para pemakai melalui web *browser*. Semua publikasi dari *website-website* tersebut dapat membentuk sebuah jaringan informasi yang sangat besar.

Halaman-halaman dari *website* akan bisa diakses melalui sebuah URL yang biasa disebut *Homepage*. URL ini mengatur halaman-halaman situs untuk menjadi sebuah hirarki, meskipun, *hyperlink-hyperlink* yang ada di halaman tersebut mengatur para pembaca dan memberitahu mereka susun keseluruhan dan bagaimana arus informasi ini berjalan. (Hidayatullah, Raden Aryadi, 2016).

2.5 HTML

HTML singkatan dari *HyperText Markup Language* yaitu sebuah Bahasa markup yang digunakan untuk membuat sebuah halaman *website*, yang menampilkan berbagai informasi dari internet dan *formatting hypertext* sederhana yang ditulis kedalam berkas format ASCII (*American Standard Code For Information Interchange*) agar dapat menghasilkan tampilan yang terintegrasi. HTML adalah bahasa yang digunakan untuk memaparkan informasi berupa *text, audio, video*, dan sebagainya.

HTML merupakan bahasa pemrograman *website* yang memiliki sintak tertentu dalam menuliskan *script* atau kode-kode, sehingga *browser* dapat menampilkan informasi dengan membaca sintak HTML. (Andrianto, P., & Nursikuwagus, A., 2017).

2.6 PHP

PHP atau kependekan dari *Hypertext Preprocessor* adalah salah satu bahasa pemrograman *open source* yang sangat cocok atau dikhususkan untuk pengembangan web dan dapat ditanamkan pada sebuah skripsi HTML. Bahasa PHP dapat dikatakan menggambarkan beberapa bahasa pemrograman seperti C, Java, dan Perl serta mudah untuk dipelajari.

PHP merupakan bahasa *scripting server – side*, dimana pemrosesan datanya dilakukan pada sisi server. Sederhananya, serverlah yang akan menerjemahkan skrip program, baru kemudian hasilnya akan dikirim kepada client yang melakukan permintaan.

Adapun pengertian lain PHP adalah akronim dari *Hypertext Preprocessor*, yaitu suatu bahasa pemrograman berbasis kode – kode (*script*) yang digunakan untuk mengolah suatu data dan mengirimkannya kembali ke web browser menjadi kode HTML. (Astria Firman, Hans F. Wowor, Xaverius Najoa, 2016).

2.7 XAMPP

XAMPP adalah *software* atau aplikasi komputer yang banyak digunakan dalam dunia *web developer* yang juga bisa dipelajari untuk membuat *website*. XAMPP adalah perangkat lunak berbasis web server yang bersifat *open source* (bebas) serta mendukung di berbagai sistem operasi seperti OS Linux, OS Windows, Mac OS, OS Windows, Mac OS, dan juga Solaris.

XAMPP bisa dilakukan untuk menghemat anggaran karena mampu menggantikan peran *web hosting* dengan cara menyimpan file ke dalam *hosting local* agar bisa dipanggil lewat *browser*. Software XAMPP dikembangkan oleh tim bernama *Apache Friends* pada tahun 2002, yang bisa didapatkan secara gratis dengan label GNU (*General Public License*). Fungsi utama XAMPP adalah sebagai server lokal untuk menyimpan berbagai jenis data *website* yang sedang dalam proses pengembangan. (Anisa Sekarningrum, 2021).

2.8 MySQL

MySQL adalah sebuah *database management system* (manajemen basis data) menggunakan perintah dasar SQL (*Structured Query Language*) yang cukup terkenal. *Database management system* (DBMS) MySQL multi pengguna dan multi alur ini sudah dipakai lebih dari 6 juta pengguna di seluruh dunia.

MySQL adalah DBMS yang *open source* dengan dua bentuk lisensi, yaitu *Free Software* (perangkat lunak bebas) dan *Shareware* (perangkat lunak berpemilik yang penggunaannya terbatas). Jadi MySQL adalah *database server* yang gratis dengan lisensi *GNU General Public License* (GPL) sehingga dapat Anda pakai untuk keperluan pribadi atau komersil tanpa harus membayar lisensi yang ada. (Budiman, Muhammad Hafizh, 2021).

2.9 Laravel

Laravel adalah salah satu *framework* PHP yang dibangun dengan konsep MVC (*Model View Controller*). Laravel adalah pengembangan *website* berbasis MVC yang ditulis dalam PHP yang dirancang untuk meningkatkan kualitas perangkat lunak dengan menyediakan sintaks yang mudah, jelas, dan menghemat waktu. Dilansir dari media online raygun.com, laravel menduduki peringkat pertama dari deretan 10 PHP *Frameworks* terbaik disusul oleh CodeIgniter, Symfony, dan lain-lain. Laravel bisa mampu mengelola *website* yang kompleks secara aman dan lebih cepat dibandingkan *framework* lain. Laravel juga menyederhanakan proses dalam pengembangan seperti *routing*, *sessions*, *caching*, dan *authentication* (Njenga, 2018).

Menurut Abdulloh (2017:3) mengatakan bahwa terdapat beberapa keunggulan yang dimiliki Laravel yaitu sebagai berikut.

1. Laravel memiliki banyak fitur yang tidak dimiliki oleh *framework* lain.
2. Laravel merupakan *framework* PHP yang ekspresif, artinya sintaks pada laravel menggunakan bahasa yang mudah dimengerti sehingga programmer pemula sekalipun akan mudah memahami kegunaan dari suatu sintaks meskipun programmer tersebut belum mempelajarinya.
3. Laravel memiliki dokumentasi yang cukup lengkap, bahkan setiap versinya memiliki dokumentasi tersendiri mulai dari cara instalasi hingga penggunaan fitur-fiturnya.
4. Laravel digunakan oleh banyak programmer sehingga banyak *library* yang mendukung Laravel yang diciptakan para programmer pecinta Laravel.
5. Laravel didukung oleh Composer sehingga *library-library* diperoleh dengan mudah dari internet menggunakan Composer.
6. Laravel memiliki *template engine* tersendiri yang diberi nama blade yang memudahkan dalam menampilkan data pada template HTML.

Fitur-fitur modern Laravel yang sangat membantu *developer* dalam membuat aplikasi adalah *Bundles*, *Eloquent ORM* (Object-Relational Mapping), *Query Builder*, *Application Logic*, *Reverse Routing*, *Resource Controller*, *Class Auto Loading*, *View Composers*, *Blade*, *IoC*, *Containers*, *Migration*, *Database Seeding*, *Unit Testing*, *Automatic Pagination*, *Form request*, *Middleware*. *Framework* laravel juga memiliki beberapa keunggulan sebagai berikut:

1. Menggunakan *Command Line Interface* (CLI) *Artisan*.
2. Menggunakan *Package manager* PHP *Composer*.
3. Penulisan kode program lebih singkat, mudah dimengerti, dan ekspresif.

2.10 Chatbot

Chatbot adalah sebuah implementasi dari bidang ilmu pengolahan bahasa alami, pembelajaran mesin, rekayasa perangkat lunak dan kecerdasan buatan. Sebuah *chatbot* akan dianggap menggunakan sebuah kecerdasan buatan, ketika masukan bahasa alami terdapat beberapa kata yang susunannya tidak sama pada sebuah kalimat, namun program mampu mempelajari bahwa masukan tersebut memiliki arti yang sama dengan masukan sebelumnya, sehingga program mampu merespon dengan balasan yang sesuai. *Chatbot* dapat dikembangkan sesuai dengan kebutuhan yang ada, namun setiap *chatbot* yang dikembangkan memiliki tipe tersendiri dalam mengeksekusi masukan pengguna. (Beatrix, 2022).

Cara *chatbot* bekerja dimulai dengan menerima input dari pengguna berupa pesan teks. Sistem kemudian menggunakan NLP untuk memproses input untuk menganalisis, mengidentifikasi, dan menafsirkan makna yang dimaksud oleh pengguna. Sistem akan memverifikasi input makna dan kondisi dari percakapan yang sedang berlangsung. Setelah sistem memahami artinya, sistem akan mencari data yang dianggap sesuai, kemudian respon berupa jawaban yang ditampilkan kepada pengguna berdasarkan struktur dan bahasa manusia. (Vincentius, Njoto dan Vioni, 2021).

2.11 Preprocessing Text

Text Preprocessing merupakan tahapan dari proses awal terhadap teks untuk mempersiapkan teks menjadi data yang akan diolah lebih lanjut. Suatu teks tidak dapat diproses langsung oleh algoritma pencarian, oleh karena itu dibutuhkan *preprocessing text* untuk mengubah teks menjadi data *numeric*. Sebuah teks yang ada harus dipisahkan, hal ini dapat dilakukan dalam beberapa tingkatan yang berbeda. Suatu dokumen atau *tweet* dapat di pecah menjadi bab, sub-bab, paragraf, kalimat dan pada akhirnya menjadi

potongkata/token. Selain itu pada tahapan ini keberadaan digit angka, huruf kapital, atau karakter-karakter yang lainnya dihilangkan dan dirubah (Ronen Feldman, 2007).

Menurut (Nugroho, 2011) Proses *Preprocessing* berfungsi untuk proses awal sebelum dokumen teks diolah pada tahap selanjutnya dimana akan dilakukan proses seleksi data yang akan di proses pada setiap dokumen. Proses ini terdiri dari beberapa proses pembersihan dokumen, yaitu *case folding*, *tokenizing*, *filtering* atau *stopword removal*, dan *stemming*.

Tahap-tahap *preprocessing text* yang digunakan adalah sebagai berikut (Rizki, Dhidik, dan Eko Suprpto. 2017) ;

1. Tokenizing

Tokenizing adalah proses memecah dokumen menjadi kumpulan kata. *Tokenization* dapat dilakukan dengan menghilangkan tanda baca dan memisahkannya per-spasi. Tahapan ini juga menghilangkan karakter-karakter tertentu seperti tanda baca dan mengubah semua token ke bentuk huruf kecil (*lower case*).

2. Stopword/Filtering

Stopwords removal merupakan proses penghilangan kata tidak penting pada deskripsi melalui pengecekan kata-kata hasil *parsing* deskripsi apakah termasuk di dalam daftar kata tidak penting (*stoplist*) atau tidak. Jika termasuk di dalam *stoplist* maka kata-kata tersebut akan di-*remove* dari deskripsi sehingga kata-kata yang tersisa di dalam deskripsi dianggap sebagai kata-kata penting atau *keywords*.

3. Stemming

Menurut (Devi Basuma, 2013) *stemming* adalah salah satu cara yang digunakan untuk meningkatkan performa sistem *information retrieval* dengan cara mentransformasi kata-kata dalam sebuah dokumenteks ke bentuk kata dasarnya untuk mengurangi ukuran *index* atau *record*, contohnya kata-kata menyukai, tersukkseskan dan disukkseskan akanditransformasi ke stem yang sama yaitu sukses.

2.12 Term Frequency Inverse Document Frequency (TF-IDF)

Term Frequency Inverse Document Frequency (TF-IDF) ialah metode pembobotan dalam bentuk integrasi antara *term frequency* dengan *inverse document frequency*. Algoritma *Term Frequency Inverse-Documents Frequency* merupakan suatu algoritma yang menggabungkan antara *Term frequency* dengan *Inverse Document Frequency*. *Term frequency* yaitu jumlah kemunculan sebuah *term* pada sebuah dokumen. *Inverse Document Frequency* yaitu pengurangan dominasi term yang sering muncul di berbagai dokumen,

dengan memperhitungkan kebalikan frekuensi dokumen yang mengandung suatu kata. Metode TF-IDF digunakan pada penelitian ini untuk memilih pertanyaan mana di *database* yang paling cocok dengan pertanyaan yang diinput oleh user dengan mencari bobot tertinggi.

Pada algoritma TF-IDF digunakan rumus untuk menghitung bobot (W) masing masing dokumen terhadap kata kunci dengan rumus yaitu :

$$W_{dt} = tf_{dt} * Id_{ft} \quad (2.1)$$

Dimana:

d = dokumen ke-d

t = kata ke-t dari kata kunci

W= bobot dokumen ke-d terhadap kata ke-t

tf = banyaknya kata yang dicari pada sebuah dokumen

IDF = *Inversed Document Frequency*

$$IDF = \log \left(\frac{n}{df} \right) \quad (2.2)$$

N = total dokumen

df = banyak dokumen yang mengandung kata yang dicari

Term weighting atau pembobotan term sangat dipengaruhi oleh hal- hal berikut ini (Mandala, 2004):

1. *Term Frequency (tf) factor*, yaitu faktor yang menentukan bobot term pada suatu dokumen berdasarkan jumlah kemunculannya dalam dokumen tersebut. Nilai jumlah kemunculan suatu kata (*term frequency*) diperhitungkan dalam pemberian bobot terhadap suatu kata. Semakin besar jumlah kemunculan suatu *term* (tf tinggi) dalam dokumen, semakin besar pula bobotnya dalam dokumen atau akan memberikan nilai kesesuaian yang semakin besar.
2. *Inverse Document Frequency (idf) factor*, yaitu pengurangan dominansi term yang sering muncul di berbagai dokumen. Hal ini diperlukan karena term yang banyak muncul di berbagai dokumen, dapat dianggap sebagai term umum (*common term*) sehingga tidak penting nilainya. Sebaliknya faktor kejarangmunculan kata (*term scarcity*) dalam koleksi dokumen harus diperhatikan dalam pemberian bobot.

2.13 Cosine Similarity

Kemiripan antar dokumen dihitung menggunakan suatu fungsi ukuran kemiripan yaitu *similarity measure*. Ukuran ini memungkinkan perbandingan dokumen sesuai dengan kemiripan relevansinya terhadap *query*. Salah satu ukuran kemiripan teks yang paling populer adalah *cosine similarity*. Ukuran ini menghitung nilai cosinus sudut antara dua vektor. Jika terdapat dua vektor dokumen d_j dan *query* q , serta t *term* diekstrak dari dari koleksi dokumen, maka nilai cosinus antara d_j dan q didefinisikan sebagai berikut (Muharromah, 2018).

Berikut adalah rumus Cosine similarity:

$$\text{Similarity} = \cos \theta = \frac{Q \cdot D}{|Q||D|} = \frac{\sum_{i=1}^n (w_{qi} \times w_{di})}{\sqrt{\sum_{i=1}^n (w_{qi})^2} \times \sqrt{\sum_{i=1}^n (w_{di})^2}} \quad (2.3)$$

Keterangan :

Q = Vektor Q, yang akan dibandingkan kemiripannya

D = Vektor D, yang akan dibandingkan kemiripannya

$Q \cdot D$ = *dot product* antara vektor Q dan vektor D

$|Q|$ = panjang vektor Q

$|D|$ = panjang vektor D

$|Q||D|$ = *cross product* antara $|Q|$ dan $|D|$

w_{qi} = bobot *term* pada *query* ke- i , = $tf \times idf$

w_{di} = bobot *term* pada dokumen ke- i istilah ke- j = $tf \times idf$

i = jumlah *term* dalam kalimat.

N = jumlah vektor.

Penulis memilih menggunakan metode *cosine similarity* dikarenakan metode ini mempunyai nilai akurasi yang tinggi dimana menurut (Ria, Victor, Hendra, dan Taslimun, 2018) kelebihan utama dari metode *cosine similarity* adalah tidak terpengaruh pada panjang pendeknya suatu dokumen. Sehingga, dengan melakukan perbandingan *keyword* yang dihasilkan, maka kedekatan antara item-pun dapat dipastikan.