

BAB V

KESIMPULAN DAN SARAN

5.1. Kesimpulan

Berdasarkan hasil eksperimen yang berupa pembentukan *dataset*, pembentukan model, pengujian model, dan perbandingan metode klasifikasi yang digunakan maka dapat ditarik kesimpulan sebagai berikut :

- a. *Dataset* karakteristik jawaban dapat dibangun dan memiliki 4 fitur. Fitur tersebut adalah nilai kesamaan kalimat, nilai kesamaan jenis kalimat, nilai kesamaan kata dan selisih jumlah kata. Semakin tinggi nilai fitur kesamaan jenis kata, kesamaan jenis kalimat dan kesamaan kalimat maka semakin tinggi pula kelas nilai yang dimiliki. Namun sebaliknya kelas nilai yang tinggi cenderung memiliki nilai fitur selisih jumlah kata yang rendah
- b. Nilai parameter optimal dalam model klasifikasi adalah C sebesar 2^5 dan γ sebesar $2^{-1}, 2^1$ dan 2^3 . Parameter tersebut dipilih sebagai nilai optimal karena memiliki nilai akurasi tertinggi yaitu 89%. Selain itu, nilai parameter tersebut memberikan nilai MAE yang kecil (0,44) dan nilai AUC yang mendekati 1 (0,931). Ini menandakan bahwa model yang dibangun dapat memprediksi data dengan nilai yang mendekati sempurna.
- c. Pada model yang dibangun, Metode SVM memiliki tingkat akurasi tertinggi yaitu 89% dengan nilai *mean absolute error* 0,044. Nilai ini lebih baik

dari metode lain yang pernah digunakan sebelumnya yaitu C4.5 dan Regresi Logistik. Namun metode Regresi Logistik memberikan nilai ROC Area yang lebih baik. Hal ini dikarenakan oleh metode SVM yang belum stabil dalam mengklasifikasikan data non-biner.

5.2. Saran

Dalam penelitian ini masih terdapat kekurangan pada jumlah data pada *dataset*. Pada penelitian selanjutnya, penulis merekomendasikan untuk menambahkan data dan fitur pada *dataset*. Fitur yang dimaksudkan seperti fitur sinonim dan hubungan antara kata dalam jawaban dan kunci jawaban. Hal ini akan meningkatkan nilai akurasi dari penilaian yang diberikan oleh sistem jika dibandingkan dengan penilaian yang diberikan oleh penilai asli. Selain itu penulis menyarankan untuk mencari metode yang dapat menstabilkan model SVM dalam mengklasifikasi data non-biner.

LAMPIRAN

1. Dataset

Tabel 1. Data pertanyaan dan jawaban yang didapatkan dari buku *Introduction of Information System*

Sumber	Data items refer to an elementary description of things, events, activities, and transactions that are recorded, classified, and stored but are not organized to convey any specific meaning. Data items can be numbers, letters, figures, sounds, or images. Examples of data items are a student grade in a class and the number of hours an employee worked in a certain week.
Pertanyaan	Apa definisi dari data?
Kunci Jawaban	Data merupakan deskripsi dasar dari sesuatu objek, kejadian, aktifitas dan transaksi yang di rekam, terklasifikasi dan disimpan tetapi belum memiliki arti yang spesifik.
A	Data merupakan suatu objek, kejadian, aktifitas dan transaksi yang direkam, diklasifikasi dan tersimpan namun belum memiliki arti yang spesifik.
B	Data merupakan deskripsi dari objek, kejadian, aktifitas dan transaksi yang direkam terklasifikasi, dan disimpan.
C	Data merupakan deskripsi dari suatu objek, kejadian, aktifitas dan transaksi.
D	Data merupakan deskripsi dasar mengenai objek dan kegiatan.

E	Data merupakan segala sesuatu yang disimpan
Sumber	Social engineering is an attack whereby the perpetrator uses social skills to trick or manipulate a legitimate employee into providing confidential company information such as passwords. The most common example of social engineering occurs when the attacker impersonates someone else on the telephone, such as a company manager or an information systems employee. The attacker says he forgot his password and asks the legitimate employee to give him a password to use. Other common exploits include posing as an exterminator, air conditioning technician, or fire marshal. Examples of social engineering abound.
Pertanyaan	Apa itu serangan menggunakan rakayasa sosial /social engineering attack?
Kunci Jawaban	Social engineering attack merupakan serangan yang dilakukan dengan menggunakan kecakapan sosial untuk menjebak atau memanipulasi seseorang agar memberikan informasi yang bersifat pribadi.
A	Social engineering attack merupakan serangan yang dilakukan seseorang dengan menggunakan kecakapan sosial untuk menjebak atau memanipulasi seseorang agar memberikan informasi yang bersifat pribadi atau rahasia.
B	Social engineering attack merupakan serangan dimana pelakunya menggunakan kecakapan sosial

	untuk mendapatkan informasi yang bersifat rahasia.
C	Social engineering attack merupakan serangan menggunakan kemampuan kecakapan sosial.
D	Social engineering attack merupakan serangan untuk mendapatkan informasi.
E	Social engineering adalah jenis serangan terhadap komputer.
Sumber	<p>Spyware is software that collects personal information about users without their consent. Two types of spyware are: keystroke loggers and screen scrapers. Keystroke loggers (also called keyloggers) record your keystrokes and your Internet Web browsing history. The purposes range from criminal (e.g., theft of passwords and sensitive personal information such as credit card numbers) to annoying (e.g., recording your Internet search history for targeted advertising). Companies have attempted to counter keyloggers by switching to other forms of input for authentication. For example, rather than typing in a password, the user has to accurately select each character in turn from a series of boxes, using a mouse. As a result, attackers have turned to screen scrapers (or screen grabbers). This software records a continuous "movie" of a screen's contents rather than simply recording keystrokes.</p>
Pertanyaan	Apa itu spyware?

<p>Kunci Jawaban</p>	<p>Spyware merupakan perangkat lunak yang mengumpulkan informasi pribadi seseorang tanpa diketahui. Tipe dari spyware diantaranya adalah keystroke loggers dan screen scarpers. Keystroke loggers merekam aktifitas penggunaan keyboard (keystroke) dan riwayat penggunaan internet pengguna. Screen scrapers merekam tampilan layar ketika pengguna menggunakan komputer.</p>
<p>A</p>	<p>Spyware merupakan perangkat lunak yang digunakan untuk mengumpulkan informasi pribadi seseorang tanpa diketahui. Tipe spyware diantaranya adalah keystroke loggers dan screen scrapers. Keystroke loggers merekam keystroke dan riwayat penggunaan internet pengguna. Sedangkan Screen scrapers merekam aktifitas layar monitor ketika pengguna menggunakan komputer.</p>
<p>B</p>	<p>Spyware merupakan perangkat lunak yang bertujuan mengumpulkan informasi pribadi seseorang tanpa dikethui. Tipe dari spyware diantaranya adalah keystroke loggers dan screen scrapers. Keystroke loggers dan screen scrapers merekam aktifitas pengguna ketika menggunakan komputer.</p>
<p>C</p>	<p>Spyware merupakan perangkat lunak yang mengumpulkan informasi pribadi seseorang tanpa diketahui. Tipe dari spyware adalah keystroke loggers dan screen scrapers.</p>

D	Spyware merupakan perangkat lunak yang mengumpulkan informasi pribadi seseorang tanpa diketahui.
E	Spyware adalah jenis perangkat lunak.
Sumber	<p>A database management system (DBMS) is a set of programs that provide users with tools to add, delete, access, and analyze data stored in one location. An organization can access the data by using query and reporting tools that are part of the DBMS or by using application programs specifically written to access the data. DBMSs also provide the mechanisms for maintaining the integrity of stored data, managing security and user access, and recovering information if the system fails. Because databases and DBMSs are essential to all areas of business, they must be carefully managed. There are a number of different database architectures, but we focus on the relational database model because it is popular and easy to use. Other database models (e.g., the hierarchical and network models) are the responsibility of the MIS function and are not used by organizational employees. Popular examples of relational databases are Microsoft Access and Oracle.</p>
Pertanyaan	Apa itu DBMS?

Kunci Jawaban	DBMS merupakan sekumpulan program yang menyediakan fungsi untuk menambah, menghapus, mengakses dan menganalisa data yang tersimpan di satu lokasi. Pengguna dapat mengakses data dengan menggunakan query dan reporting tools yang merupakan bagian dari DBMS atau menggunakan program aplikasi yang secara spesifik dibuat untuk mengakses data.
A	DBMS merupakan sekumpulan program yang menyediakan fungsi untuk menambah, menghapus, dan menganalisa data yang telah tersimpan. Pengguna mengakses data dengan menggunakan query dan reporting tools atau menggunakan program aplikasi yang secara spesifik digunakan untuk mengakses data.
B	DBMS merupakan sekumpulan program yang menyediakan fungsi untuk menambah, menghapus, mengakses dan menganalisa data yang tersimpan di satu lokasi. Pengguna dapat mengakses data dengan menggunakan query atau reporting tools.
C	DBMS merupakan program yang menyediakan fungsi untuk menganalisa data yang telah tersimpan di suatu lokasi.
D	Dengan DBMS pengguna dapat menghapus, menambah, mengakses dan menganalisa data yang tersimpan.
E	DBMS merupakan sekumpulan program.
Sumber	Systems software is a set of instructions that serves primarily as an intermediary between computer hardware and application programs.

	Systems software provides important self-regulatory functions for computer systems, such as loading itself when the computer is first turned on and providing commonly used sets of instructions for all applications. <i>Systems programming</i> refers to both the creation and maintenance of systems software.
Pertanyaan	Apa itu perangkat lunak sistem / system software?
Kunci Jawaban	System software merupakan sekumpulan instruksi yang memiliki fungsi utama sebagai penghubung antara perangkat keras komputer dengan program aplikasi.
A	System software adalah sekumpulan instruksi yang memiliki fungsi utama sebagai penghubung antara perangkat keras komputer dengan program aplikasi.
B	System software merupakan instruksi yang memiliki fungsi sebagai perantara antara perangkat keras komputer dengan program aplikasi.
C	System software merupakan sekumpulan instruksi yang bertugas sebagai perangkat lunak penghubung.
D	System software merupakan sekumpulan instruksi pada komputer.
E	System software merupakan perangkat lunak.
Sumber	open-source software is software whose source code is available at no cost to developers or users. Open-source software is not shareware

	<p>or freeware. <i>Shareware</i> typically allows no access to the underlying source code. <i>Freeware</i> is copyrighted software that is made available to the user free of charge for an unlimited time. In contrast, <i>opensource</i> software is copyrighted and distributed with license terms ensuring that the source code will always be available.</p>
Pertanyaan	Apa definisi dari perangkat lunak open-source?
Kunci Jawaban	Perangkat lunak open-source merupakan perangkat lunak yang membagikan kodenya secara gratis kepada pembuat aplikasi.
A	Perangkat lunak open-source merupakan perangkat lunak yang membagikan kodenya dengan gratis kepada pembuat perangkat lunak.
B	Perangkat lunak open-source merupakan perangkat lunak yang memiliki kode yang dapat digunakan secara gratis.
C	Perangkat lunak open-source menyediakan kode sumber secara gratis untuk pembuat aplikasi.
D	Merupakan perangkat lunak yang dibagikan dengan gratis.
E	Merupakan perangkat lunak gratis.
Pertanyaan	Mengapa perangkat lunak open-source berbeda dengan shareware?
Kunci Jawaban	Perangkat lunak open-source menyediakan kode programnya kepada pembuat aplikasi, sedangkan shareware pada umumnya tidak memberikan akses ke kode programnya.

A	Perangkat lunak open-source memberikan kode programnya kepada pembuat aplikasi, sedangkan shareware tidak memberikan akses ke kode programnya.
B	Karena perangkat lunak open-source menyediakan kode programnya ke pembuat aplikasi sedangkan shareware tidak.
C	Karena perangkat lunak open-source menyediakan kode programnya.
D	Karena open-source menyediakan programnya gratis.
E	Karena open-source gratis.
Sumber	Peer-to-peer (P2P) processing is a type of client/server distributed processing in which each computer acts as <i>both</i> a client and a server. Each computer can access (as assigned for security or integrity purposes) all files on all other computers. There are three basic types of peer-to-peer processing. The first accesses unused CPU power among networked computers. A well-known application of this type is SETI@home (http://setiathome.ssl.berkeley.edu) (Figure TG4.10). These applications are from open-source projects and can be downloaded at no cost. The second form of peer-to-peer is real-time, person-to-person collaboration, such as America Online's Instant Messenger. Companies such as Groove Networks (www.groove.net) have introduced P2P collaborative applications that

	<p>use buddy lists to establish a connection and allow real-time collaboration within the application. The third peer-to-peer category is advanced search and file sharing. This category is characterized by natural language searches of millions of peer systems and lets users discover other users, not just data and Web pages. One example of this is BitTorrent.</p>
Pertanyaan	Apa itu proses peer-to-peer?
Kunci Jawaban	Peer-to-peer merupakan tipe jaringan client/server dimana setiap komputer yang terhubung bertindak sebagai client dan sekaligus server. Setiap komputer yang terhubung dapat mengakses data yang terdapat pada komputer lain dalam satu jaringan yang sama.
A	Peer-to-peer merupakan tipe jaringan yang memungkinkan setiap komputer yang terhubung bertindak sebagai client dan sekaligus server. Oleh karena itu, setiap komputer dapat mengakses data yang terdapat di komputer lain dalam jaringan yang sama.
B	Dalam jaringan peer-to-peer, setiap komputer bertindak sebagai client sekaligus server. Komputer yang terhubung dapat mengakses data pada komputer lain.
C	Peer-to-peer merupakan jaringan yang menyebabkan setiap komputer bertindak sebagai client sekaligus server.
D	Peer-to-peer merupakan jenis jaringan client/server.

E	Merupakan tipe jaringan.
Sumber	Natural language processing (NLP) refers to communicating with a computer in the user's native language. To understand a natural language inquiry, a computer must have the knowledge to analyze and then interpret the input. This knowledge may include linguistic knowledge about words, domain knowledge (knowledge of a narrowly defined, specific area, such as student registration or air travel), commonsense knowledge, and even knowledge about the users and their goals. Once the computer understands the input, it can perform the desired action.
Pertanyaan	Bagaimana komputer dapat mengerti input dalam bahasa natural?
Kunci Jawaban	Untuk mengerti input dalam bahasa natural, komputer harus memiliki pengetahuan untuk menganalisa dan menginterpretasikan input yang diberikan. Pengetahuan ini meliputi pengetahuan linguistik mengenai kata, pengetahuan domain, pengetahuan mengenai rasionalitas, dan bahkan pengetahuan mengenai pengguna dan tujuannya.
A	Untuk mengerti input dalam bahasa natural, komputer harus memiliki pengetahuan untuk menganalisa dan menginterpretasikan input dari pengguna. Pengetahuan ini meliputi pengetahuan linguistik mengenai kata, pengetahuan domain

	dan pengetahuan mengenai pengguna dan tujuannya.
B	Untuk mengerti input dalam bahasa natural, komputer harus memiliki pengetahuan terhadap input. Pengetahuan ini meliputi pengetahuan linguistik mengenai kata dan pengetahuan mengenai pengguna dan tujuannya.
C	Untuk mengerti input dengan bahasa natural, komputer harus memiliki pengetahuan terhadap input. Pengetahuan ini adalah pengetahuan linguistik mengenai kata.
D	Untuk mengerti input dalam bahasa natural, komputer harus memiliki pengetahuan terhadap input.
E	Komputer menggunakan pengetahuan untuk mengerti input dengan bahasa natural.
Sumber	Fuzzy logic is a branch of mathematics that deals with uncertainties by simulating the process of human reasoning. The rationale behind fuzzy logic is that decision making is not always a matter of black and white, true or false. It often involves gray areas where the term <i>maybe</i> is more appropriate. A computer programmed to use fuzzy logic precisely handles subjective concepts that humans do not define precisely. A term such as "warm" is related, via precisely defined formulas, to an imprecise concept. For example, where the concept is "income," "high" could have values ranging over \$200,000 per year and "moderate"

	could have values ranging from \$75,000 to \$150,000 per year. A loan officer at a bank might use fuzzy values such as high and moderate when considering a loan application.
Pertanyaan	Apa itu logika fuzzy?
Kunci Jawaban	Logika fuzzy merupakan cabang dari ilmu matematika yang mempelajari tentang ketidakpastian dengan mensimulasikan proses human reasoning. Logika fuzzy dapat dengan tepat menanggulangi konsep subjektif yang tidak bisa ditentukan oleh manusia secara tepat.
A	Logika fuzzy merupakan ilmu yang mempelajari tentang ketidakpastian dengan mensimulasikan proses human reasoning. Logika fuzzy dapat dengan tepat menanggulangi konsep subjektif yang tidak bisa ditentukan oleh manusia secara tepat.
B	Logika fuzzy merupakan cabang ilmu matematika yang mempelajari tentang ketidakpastian dengan mensimulasikan proses human reasoning.
C	Logika fuzzy merupakan cabang ilmu yang dapat menanggulangi hal yang tidak bisa ditentukan oleh manusia secara tepat.
D	Logika fuzzy merupakan ilmu matematika mengenai ketidakpastian.
E	Logika fuzzy merupakan konsep matematika.
Sumber	DDOS attack typically by using malicious software. These computers are called <i>zombies</i> , or <i>bots</i> . The attacker uses these bots (which

	form a <i>botnet</i>) to deliver a coordinated stream of information requests to a target computer, causing it to crash.
Pertanyaan	Apa itu serangan DDOS?
Kunci Jawaban	Serangan DDOS merupakan serangan yang biasanya menggunakan perangkat lunak jahat untuk menginfeksi banyak komputer. Komputer yang terinfeksi biasanya disebut zombies atau bots. Pengguna DDOS menggunakan bots untuk mengirimkan arus data dengan jumlah yang besar ke komputer target yang dapat menyebabkan gangguan sistem.
A	Serangan DDOS merupakan serangan yang menggunakan banyak komputer yang terinfeksi oleh perangkat lunak jahat. Komputer yang terinfeksi biasanya disebut zombies atau bots. Pengguna DDOS menggunakan bots untuk mengirimkan aliran data yang besar ke komputer target sehingga menyebabkan gangguan terhadap sistem.
B	Serangan DDOS merupakan serangan yang biasanya menggunakan perangkat lunak jahat untuk meninfeksi banyak komputer. Komputer yang terinfeksi disebut bots. bots digunakan untuk mengirimkan aliran data yang besar ke komputer target.
C	Serangan DDOS merupakan serangan yang menggunakan komputer yang terinfeksi untuk mengirimkan aliran data yang besar ke komputer target. Komputer yang terinfeksi disebut bots.

D	Serangan DDOS merupakan serangan yang menggunakan perangkat lunak jahat.
E	Serangan DDOS menggunakan komputer zombie.

Tabel 2. Data pertanyaan dan jawaban yang didapatkan dari hasil wawancara

Jelaskan definisi dari hak kepemilikan perangkat lunak software (software ownership right)	
Hak kepemilikan perangkat lunak merupakan pembatasan yang diberlakukan kepada penggunaan, penyalinan, dan modifikasi yang telah ditetapkan untuk pemegang hak dari perangkat lunak. Hak kepemilikan perangkat lunak menjelaskan aktivitas apa saja yang diperbolehkan dalam menggunakan perangkat lunak tersebut dalam bentuk lisensi. Hal ini menyebabkan pemilik hak tidak membeli sepenuhnya. Hak kepemilikan perangkat lunak dibuat setelah perangkat lunak atau program tersebut dikembangkan.	A
Hak kepemilikan perangkat lunak adalah batasan yang diberlakukan kepada penggunaan, penyalinan dan modifikasi yang telah ditetapkan untuk pemilik hak perangkat lunak. Hak kepemilikan perangkat lunak menjelaskan aktivitas apa saja yang diperbolehkan dalam menggunakan perangkat lunak tersebut yang berbentuk lisensi. Hal ini menyebabkan pemilik hak tidak membeli sepenuhnya. Hak kepemilikan akan dibuat setelah perangkat lunak atau program dikembangkan.	A
Hak kepemilikan perangkat lunak adalah pembatasan pada pemberlakuan penggunaan, penyalinan, dan	A

<p>modifikasi yang telah ditetapkan untuk pemegang hak dari perangkat lunak. Hak kepemilikan perangkat lunak menerangkan mengenai segala aktivitas yang diperbolehkan dalam penggunaan perangkat lunak yang berbentuk lisensi. Hal ini menyebabkan pemilik hak tidak membeli sepenuhnya. Hak kepemilikan perangkat lunak dibuat setelah perangkat lunak dikembangkan.</p>	
<p>Hak kepemilikan perangkat lunak merupakan pembatasan yang diberlakukan kepada penggunaan, penyalinan, dan modifikasi yang telah ditetapkan untuk pemegang hak dari perangkat lunak. Hak kepemilikan perangkat lunak menjelaskan aktivitas apa saja yang diperbolehkan dalam menggunakan perangkat lunak tersebut dalam bentuk lisensi. Hal ini menyebabkan pemilik hak tidak membeli sepenuhnya.</p>	A
<p>Hak kepemilikan perangkat lunak merupakan pembatasan yang diberlakukan kepada penggunaan, penyalinan, dan modifikasi yang telah ditetapkan untuk pemegang hak dari perangkat lunak. Hak kepemilikan perangkat lunak menjelaskan aktivitas apa saja yang diperbolehkan dalam menggunakan perangkat lunak tersebut dalam bentuk lisensi.</p>	B
<p>Hak kepemilikan perangkat lunak merupakan pembatasan yang diberlakukan kepada penggunaan, penyalinan, dan modifikasi yang telah ditetapkan untuk pemegang hak dari perangkat lunak. Hak kepemilikan perangkat lunak dibuat setelah perangkat lunak atau program tersebut dikembangkan.</p>	B

Hak kepemilikan perangkat lunak menjelaskan aktivitas apa saja yang diperbolehkan dalam menggunakan perangkat lunak dalam bentuk lisensi. Hal ini menyebabkan pemilik hak tidak membeli sepenuhnya. Hak kepemilikan perangkat lunak dibuat setelah perangkat lunak atau program tersebut dikembangkan.	B
Hak kepemilikan perangkat lunak merupakan pembatasan yang diberlakukan kepada penggunaan, penyalinan, dan modifikasi yang telah ditetapkan untuk pemegang hak dari perangkat lunak.	C
Hak kepemilikan perangkat lunak menjelaskan aktivitas apa yang diperbolehkan dalam menggunakan perangkat lunak. Ini menyebabkan pemilik hak tidak membeli sepenuhnya.	C
Hak kepemilikan perangkat lunak menyebabkan pemilik tidak membeli program sepenuhnya. Hak kepemilikan perangkat lunak dibuat setelah perangkat lunak atau program tersebut dikembangkan.	C
Hak kepemilikan perangkat lunak merupakan pembatasan penggunaan perangkat lunak oleh pemegang hak.	D
Hak kepemilikan perangkat lunak adalah batasan penggunaan perangkat lunak oleh pemegang hak perangkat lunak.	D
Hak kepemilikan perangkat lunak merupakan batasan yang diberikan dalam penggunaan perangkat lunak oleh pemegang hak	D
Hak yang dimiliki perusahaan yang berisi perangkat lunak	E

Hak yang berisi instruksi penggunaan perangkat lunak perusahaan	E
Hak ini adalah segala sesuatu yang bisa dilakukan kepada perangkat lunak	E
Jelaskan apa yang disebut dengan functional area information system	
Functional area information systems adalah perangkat lunak yang diimplementasikan untuk mendukung functional area yang spesifik. Functional area information systems bertujuan untuk meningkatkan keefektifan dan keefisienan pada suatu functional area. Functional area yang dimaksud adalah accounting, finance, marketing, operations, dan human resource management.	A
Functional area information systems merupakan pengimplementasian perangkat lunak untuk mendukung functional area yang spesifik. Functional area information systems bertujuan untuk meningkatkan keefektifan dan keefisienan suatu functional area. Functional area tersebut adalah accounting, finance, marketing, operations, dan human resource management.	A
Functional area information systems ialah perangkat lunak yang diimplementasikan guna mendukung functional area yang spesifik. Functional area information systems bertujuan agar meningkatnya keefektifan dan keefisienan suatu functional area. Functional area tersebut diantaranya adalah finance, accounting, marketing, operations, dan human resource management.	A

<p>Functional area information systems adalah pengimplementasian perangkat lunak untuk mendukung functional area yang spesifik. Functional area yang dimaksud adalah accounting, finance, marketing, operations, dan human resource management. Functional area information systems sendiri bertujuan untuk meningkatkan keefektifan dan keefisienan pada suatu functional area.</p>	A
<p>Functional area information systems adalah perangkat lunak yang diimplementasikan untuk mendukung functional area yang spesifik. Functional area yang dimaksud adalah accounting, finance, marketing, operations, dan human resource management.</p>	B
<p>Functional area information systems ialah perangkat lunak yang diimplementasikan guna mendukung functional area yang spesifik. Functional area tersebut diantaranya adalah finance, accounting, marketing, operations, dan human resource management.</p>	B
<p>Functional area information systems adalah perangkat lunak yang diimplementasikan untuk mendukung functional area yang spesifik. Functional area information systems bertujuan untuk meningkatkan kerja dalam functional area. Functional area tersebut adalah accounting, finance, dan marketing.</p>	B
<p>Functional area information systems merupakan pengimplementasian perangkat lunak untuk mendukung functional area yang spesifik.</p>	B

Functional area tersebut diantaranya adalah marketing, finance, dan human resource management.	
Functional area information systems adalah perangkat lunak yang diimplementasikan untuk mendukung functional area. Functional area information systems bertujuan untuk meningkatkan kinerja pada satu functional area.	C
Functional area information systems ialah perangkat lunak yang diimplementasikan untuk mendukung functional area yang spesifik. Functional area information systems bertujuan untuk mengoptimalkan functional area.	C
Functional area information systems adalah perangkat lunak yang digunakan untuk mendukung functional area.	D
Functional area information systems adalah perangkat lunak yang bertujuan untuk meningkatkan efektifitas dan efisiensi dari sebuah functional area.	D
Functional area information systems merupakan perangkat lunak untuk meningkatkan kinerja pada accounting dan human resource.	D
Functional area information systems adalah software dalam perusahaan.	E
Functional area information systems adalah sekumpulan software di perusahaan.	E
Functional area information systems adalah sistem di perusahaan.	E
Apa tujuan dari perancangan IMK?	

Tujuan dari perancangan IMK adalah agar dapat membangun atau memperbaiki safety, utility, effectiveness, efficiency, dan usability sistem yang berkaitan dengan komputer.	A
Tujuan dari perancangan IMK adalah agar dapat membangun safety, utility, effectiveness, efficiency, dan usability sistem komputer.	A
Tujuan dari perancangan IMK adalah agar terbangunnya sistem komputer dengan safety, utility, effectiveness, efficiency, dan usability yang lebih baik.	A
Tujuan dari perancangan IMK adalah agar dapat membangun atau memperbaiki safety, utility, effectiveness, efficiency sistem komputer.	A
Tujuan dari perancangan IMK adalah agar dapat membangun safety, utility, effectiveness sistem komputer.	B
Tujuan dari perancangan IMK adalah untuk membangun atau memperbaiki safety, utility dan usability sistem komputer.	B
Tujuan dari perancangan IMK adalah membangun atau memperbaiki safety, utility, dan effectiveness sistem komputer.	B
Tujuan dari perancangan IMK adalah agar dapat membangun safety, utility dan usability.	C
Tujuan dari perancangan IMK adalah membangun dan memperbaiki safety dan usability komputer.	C
Tujuan dari perancangan IMK adalah agar dapat membangun efficiency dan usability dari komputer.	C
Tujuan dari perancangan IMK adalah membangun utility komputer.	D

Tujuan dari perancangan IMK adalah meningkatkan safety komputer.	D
Tujuan dari perancangan IMK adalah memberikan utility komputer.	D
Perancangan IMK bertujuan memperbaiki komputer.	E
Perancangan IMK bertujuan untuk memberikan safety komputer.	E
Perancangan IMK memiliki tujuan memudahkan penggunaan komputer.	E

Tabel 3. *Dataset yang terbentuk setelah proses Natural Language Processing*

Selisih Kata	Nilai Kesamaan Jenis Kata	Nilai Kesamaan Jenis Kalimat	Nilai Kesamaan Kalimat	Kelas Nilai
2	0.7308	0.8587	0.8333	A
0	0.7692	0.9271	0.7895	A
8	0.913	0.8214	0.8824	A
0	0.9091	0.8889	0.9091	A
1	0.9091	0.8889	0.9524	A
0	0.9091	1	1	A
1	0.8462	0.8462	0.9167	A
1	0.7143	0.8462	0.9167	A
2	0.8462	0.8462	0.8333	A
2	0.5	0.8462	0.8462	A
2	0.9167	0.9167	0.9333	A
1	0.8148	0.8911	0.8261	A
4	0.8421	0.706	0.8333	A

1	0.8571	0.8571	0.8	A
1	0.9	0.9	0.8571	A
3	0.8	0.6778	1	A
2	0.8333	0.8291	0.8	A
2	0.875	0.9091	0.875	A
1	0.8571	0.9444	0.8333	A
1	0.9091	0.9167	0.8889	A
12	0.6957	0.5714	0.7647	B
13	0.6087	0.5714	0.7059	B
13	0.6957	0.8056	0.6471	B
10	0.8095	0.8	0.8095	B
11	0.8095	0.8	0.7619	B
6	0.6818	0.7609	0.6364	B
12	0.6364	0.7259	0.6364	B
3	0.6923	0.6923	0.75	B
3	0.7692	0.7692	0.75	B
3	0.7692	0.7692	0.75	B
5	0.6154	0.6154	0.5833	B
4	0.5714	0.6923	0.6923	B
3	0.5714	0.5714	0.6667	B
5	0.6296	0.582	0.6667	B
7	0.7895	0.6731	0.7778	B
3	0.7143	0.7143	0.7857	B
0	0.5	0.6364	0.75	B
6	0.625	0.3884	0.8182	B
4	0.7647	0.7424	0.6667	B
6	0.6875	0.5625	0.75	B

4	0.5909	0.5093	0.6667	B
3	0.6923	0.375	0.8889	B
21	0.5217	0.3889	0.5882	C
23	0.4583	0.5557	0.4118	C
18	0.4348	0.4679	0.4706	C
12	0.5	0.5871	0.4545	C
12	0.5	0.506	0.4545	C
5	0.5385	0.5385	0.5833	C
5	0.5385	0.5385	0.5833	C
6	0.4615	0.4615	0.5	C
7	0.4615	0.4615	0.4615	C
7	0.5	0.5	0.5	C
14	0.48	0.4286	0.5909	C
15	0.4444	0.3889	0.5294	C
5	0.4	0.4	0.4375	C
1	0.5385	0.4615	0.6667	C
10	0.5333	0.3561	0.6364	C
7	0.5882	0.4545	0.5333	C
11	0.4375	0.2262	0.4375	C
8	0.4762	0.5	0.5294	C
5	0.4615	0.4615	0.6667	C
24	0.3478	0.3571	0.4118	D
23	0.2917	0.3571	0.3529	D
24	0.2917	0.3571	0.4118	D
22	0.381	0.4682	0.3333	D
18	0.3913	0.3379	0.3333	D
19	0.4545	0.2405	0.4348	D

6	0.3571	0.3571	0.3846	D
7	0.3846	0.3846	0.4167	D
6	0.2857	0.2857	0.3333	D
7	0.4167	0.4167	0.4	D
20	0.24	0.2727	0.3636	D
18	0.35	0.1083	0.3529	D
8	0.3846	0.3846	0.4286	D
5	0.25	0.0714	0.5714	D
11	0.4	0.1864	0.4167	D
9	0.4706	0.4091	0.4	D
11	0.3125	0.3273	0.3529	D
15	0.3	0.2708	0.375	D
7	0.3077	0.2143	0.3	D
28	0.1154	0.098	0.15	E
28	0.1538	0.0906	0.15	E
30	0.1304	0.1536	0.2353	E
25	0.1739	0.265	0.1739	E
24	0.1667	0.2454	0.1667	E
25	0.1739	0.265	0.1739	E
8	0.2	0.125	0.3333	E
7	0.2	0.125	0.4167	E
6	0.1111	0.1111	0.2857	E
10	0.1538	0.1538	0.2308	E
8	0.2308	0.2308	0.1765	E
24	0.1154	0.1349	0.1304	E
20	0.1667	0.1607	0.2353	E
9	0.2143	0.2143	0.2667	E

5	0.2727	0	0.5714	E
14	0.1333	0	0.2727	E
12	0.2778	0.1429	0.3571	E
13	0.25	0.2361	0.3125	E
18	0.1429	0.0905	0.1765	E
9	0.1538	0	0.1	E

Tabel 4. Hasil normalisasi *dataset*.

Selisih Kata	Nilai Kesamaan Jenis Kata	Nilai Kesamaan Jenis Kalimat	Nilai Kesamaan Kalimat	Kelas Nilai
0.0667	0.7308	0.8587	0.8333	A
0.0000	0.7692	0.9271	0.7895	A
0.2667	0.9130	0.8214	0.8824	A
0.0000	0.9091	0.8889	0.9091	A
0.0333	0.9091	0.8889	0.9524	A
0.0000	0.9091	1.0000	1.0000	A
0.0333	0.8462	0.8462	0.9167	A
0.0333	0.7143	0.8462	0.9167	A
0.0667	0.8462	0.8462	0.8333	A
0.0667	0.5000	0.8462	0.8462	A
0.0667	0.9167	0.9167	0.9333	A
0.0333	0.8148	0.8911	0.8261	A
0.1333	0.8421	0.7060	0.8333	A
0.0333	0.8571	0.8571	0.8000	A
0.0333	0.9000	0.9000	0.8571	A
0.1000	0.8000	0.6778	1.0000	A
0.0667	0.8333	0.8291	0.8000	A
0.0667	0.8750	0.9091	0.8750	A
0.0333	0.8571	0.9444	0.8333	A
0.0333	0.9091	0.9167	0.8889	A
0.4000	0.6957	0.5714	0.7647	B
0.4333	0.6087	0.5714	0.7059	B
0.4333	0.6957	0.8056	0.6471	B

0.3333	0.8095	0.8000	0.8095	B
0.3667	0.8095	0.8000	0.7619	B
0.2000	0.6818	0.7609	0.6364	B
0.4000	0.6364	0.7259	0.6364	B
0.1000	0.6923	0.6923	0.7500	B
0.1000	0.7692	0.7692	0.7500	B
0.1000	0.7692	0.7692	0.7500	B
0.1667	0.6154	0.6154	0.5833	B
0.1333	0.5714	0.6923	0.6923	B
0.1000	0.5714	0.5714	0.6667	B
0.1667	0.6296	0.5820	0.6667	B
0.2333	0.7895	0.6731	0.7778	B
0.1000	0.7143	0.7143	0.7857	B
0.0000	0.5000	0.6364	0.7500	B
0.2000	0.6250	0.3884	0.8182	B
0.1333	0.7647	0.7424	0.6667	B
0.2000	0.6875	0.5625	0.7500	B
0.1333	0.5909	0.5093	0.6667	B
0.1000	0.6923	0.3750	0.8889	B
0.7000	0.5217	0.3889	0.5882	C
0.7667	0.4583	0.5557	0.4118	C
0.6000	0.4348	0.4679	0.4706	C
0.4000	0.5000	0.5871	0.4545	C
0.4000	0.5000	0.5060	0.4545	C
0.1667	0.5385	0.5385	0.5833	C
0.1667	0.5385	0.5385	0.5833	C
0.2000	0.4615	0.4615	0.5000	C
0.2333	0.4615	0.4615	0.4615	C
0.2333	0.5000	0.5000	0.5000	C
0.4667	0.4800	0.4286	0.5909	C
0.5000	0.4444	0.3889	0.5294	C
0.1667	0.4000	0.4000	0.4375	C
0.0333	0.5385	0.4615	0.6667	C
0.3333	0.5333	0.3561	0.6364	C
0.2333	0.5882	0.4545	0.5333	C
0.3667	0.4375	0.2262	0.4375	C
0.2667	0.4762	0.5000	0.5294	C
0.1667	0.4615	0.4615	0.6667	C
0.8000	0.3478	0.3571	0.4118	D

0.7667	0.2917	0.3571	0.3529	D
0.8000	0.2917	0.3571	0.4118	D
0.7333	0.3810	0.4682	0.3333	D
0.6000	0.3913	0.3379	0.3333	D
0.6333	0.4545	0.2405	0.4348	D
0.2000	0.3571	0.3571	0.3846	D
0.2333	0.3846	0.3846	0.4167	D
0.2000	0.2857	0.2857	0.3333	D
0.2333	0.4167	0.4167	0.4000	D
0.6667	0.2400	0.2727	0.3636	D
0.6000	0.3500	0.1083	0.3529	D
0.2667	0.3846	0.3846	0.4286	D
0.1667	0.2500	0.0714	0.5714	D
0.3667	0.4000	0.1864	0.4167	D
0.3000	0.4706	0.4091	0.4000	D
0.3667	0.3125	0.3273	0.3529	D
0.5000	0.3000	0.2708	0.3750	D
0.2333	0.3077	0.2143	0.3000	D
0.9333	0.1154	0.0980	0.1500	E
0.9333	0.1538	0.0906	0.1500	E
1.0000	0.1304	0.1536	0.2353	E
0.8333	0.1739	0.2650	0.1739	E
0.8000	0.1667	0.2454	0.1667	E
0.8333	0.1739	0.2650	0.1739	E
0.2667	0.2000	0.1250	0.3333	E
0.2333	0.2000	0.1250	0.4167	E
0.2000	0.1111	0.1111	0.2857	E
0.3333	0.1538	0.1538	0.2308	E
0.2667	0.2308	0.2308	0.1765	E
0.8000	0.1154	0.1349	0.1304	E
0.6667	0.1667	0.1607	0.2353	E
0.3000	0.2143	0.2143	0.2667	E
0.1667	0.2727	0.0000	0.5714	E
0.4667	0.1333	0.0000	0.2727	E
0.4000	0.2778	0.1429	0.3571	E
0.4333	0.2500	0.2361	0.3125	E
0.6000	0.1429	0.0905	0.1765	E
0.3000	0.1538	0.0000	0.1000	E

2. Hasil yang Diberikan Oleh Perangkat Lunak WEKA

Gambar 1. Performa klasifikasi oleh SVM kernel linier dengan nilai $C = 2^5$

```
=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances      89           89    %
Incorrectly Classified Instances    11           11    %
Kappa statistic                    0.8623
Mean absolute error                 0.044
Root mean squared error             0.2098
Relative absolute error             13.7536 %
Root relative squared error         52.4354 %
Total Number of Instances          100

=== Detailed Accuracy By Class ===

      TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class
      0.9      0       1          0.9    0.947     0.95     A
      1      0.038   0.88       1      0.936     0.981     B
      0.842   0.012   0.941     0.842  0.889     0.915     C
      0.789   0.049   0.789     0.789  0.789     0.87      D
      0.9     0.038   0.857     0.9    0.878     0.931     E
Weighted Avg. 0.89    0.028   0.894     0.89   0.89      0.931

=== Confusion Matrix ===

 a b c d e <-- classified as
18 2 0 0 0 | a = A
 0 22 0 0 0 | b = B
 0 1 16 2 0 | c = C
 0 0 1 15 3 | d = D
 0 0 0 2 18 | e = E
```

Gambar 2. Performa klasifikasi oleh SVM kernel RBF dengan nilai $C = 2^5$ dan $\text{Gamma} = 2^{-1}$

```

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances      89           89      %
Incorrectly Classified Instances    11           11      %
Kappa statistic                    0.8623
Mean absolute error                 0.044
Root mean squared error             0.2098
Relative absolute error             13.7536 %
Root relative squared error         52.4354 %
Total Number of Instances          100

=== Detailed Accuracy By Class ===

```

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.9	0	1	0.9	0.947	0.95	A
	1	0.038	0.88	1	0.936	0.981	B
	0.842	0.012	0.941	0.842	0.889	0.915	C
	0.789	0.049	0.789	0.789	0.789	0.87	D
	0.9	0.038	0.857	0.9	0.878	0.931	E
Weighted Avg.	0.89	0.028	0.894	0.89	0.89	0.931	

```

=== Confusion Matrix ===

```

a	b	c	d	e	<-- classified as
18	2	0	0	0	a = A
0	22	0	0	0	b = B
0	1	16	2	0	c = C
0	0	1	15	3	d = D
0	0	0	2	18	e = E

Gambar 3. Performa klasifikasi oleh SVM kernel RBF dengan nilai $C = 2^5$ dan $\text{Gamma} = 2^1$

```

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances      89           89      %
Incorrectly Classified Instances    11           11      %
Kappa statistic                    0.8623
Mean absolute error                 0.044
Root mean squared error            0.2098
Relative absolute error             13.7536 %
Root relative squared error        52.4354 %
Total Number of Instances          100

=== Detailed Accuracy By Class ===

          TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class
          0.9     0.013   0.947     0.9     0.923     0.944     A
          0.955   0.038   0.875     0.955   0.913     0.958     B
          0.842   0.025   0.889     0.842   0.865     0.909     C
          0.789   0.037   0.833     0.789   0.811     0.876     D
          0.95    0.025   0.905     0.95    0.927     0.963     E
Weighted Avg.  0.89    0.028   0.89      0.89    0.889     0.931

=== Confusion Matrix ===

 a  b  c  d  e  <-- classified as
18  2  0  0  0 | a = A
 1 21  0  0  0 | b = B
 0  1 16  2  0 | c = C
 0  0  2 15  2 | d = D
 0  0  0  1 19 | e = E

```

Gambar 4. Performa klasifikasi oleh SVM kernel RBF dengan nilai $C = 2^5$ dan $\text{Gamma} = 2^3$

```

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances      89           89    %
Incorrectly Classified Instances    11           11    %
Kappa statistic                    0.8623
Mean absolute error                 0.044
Root mean squared error            0.2098
Relative absolute error            13.7536 %
Root relative squared error        52.4354 %
Total Number of Instances         100

=== Detailed Accuracy By Class ===

          TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class
          0.9      0        1          0.9    0.947     0.95     A
          1        0.038   0.88       1      0.936     0.981     B
          0.842    0.025   0.889     0.842  0.865     0.909     C
          0.789    0.049   0.789     0.789  0.789     0.87     D
          0.9      0.025   0.9        0.9    0.9       0.938     E
Weighted Avg.  0.89    0.028   0.892     0.89   0.89     0.931

=== Confusion Matrix ===

 a b c d e <-- classified as
18 2 0 0 0 | a = A
 0 22 0 0 0 | b = B
 0 1 16 2 0 | c = C
 0 0 2 15 2 | d = D
 0 0 0 2 18 | e = E

```

3. Implementasi Algoritma

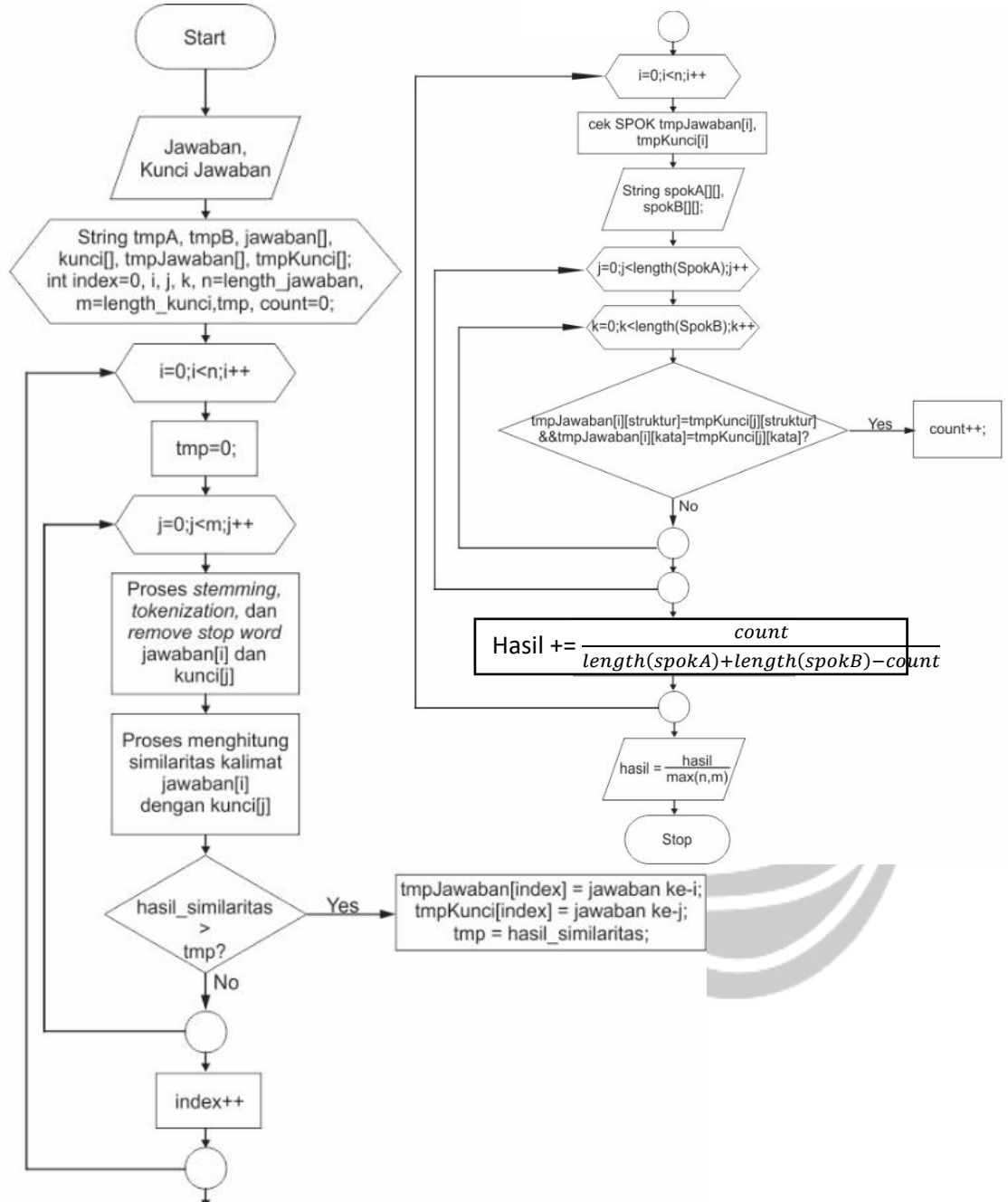
Tabel 5 Studi kasus soal dan jawaban

Pertanyaan	Jelaskan definisi dari hak kepemilikan perangkat lunak (software ownership right)
Kunci Jawaban	Hak kepemilikan perangkat lunak merupakan pembatasan yang diberlakukan kepada penggunaan, penyalinan, dan modifikasi yang telah ditetapkan untuk pemegang hak dari perangkat lunak. Hak kepemilikan perangkat lunak menjelaskan aktivitas apa saja yang diperbolehkan dalam menggunakan perangkat lunak tersebut dalam bentuk lisensi. Hal ini menyebabkan pemilik hak tidak membeli sepenuhnya. Hak kepemilikan perangkat lunak dibuat setelah perangkat lunak atau program tersebut dikembangkan.
Jawaban 1	Hak kepemilikan perangkat lunak merupakan pembatasan yang diberlakukan kepada penggunaan, penyalinan, dan modifikasi yang telah ditetapkan untuk pemegang hak dari perangkat lunak. Hak kepemilikan perangkat lunak menjelaskan aktivitas apa saja yang diperbolehkan dalam menggunakan

	perangkat lunak tersebut dalam bentuk lisensi. Hal ini menyebabkan pemilik hak tidak membeli sepenuhnya
Jawaban 2	Hak kepemilikan perangkat lunak merupakan pembatasan penggunaan perangkat lunak oleh pemegang hak.



3.1. Proses Perhitungan Nilai Kesamaan Struktur Kalimat (SPOK)



Gambar 5. Flow chart proses perhitungan nilai kesamaan struktur kalimat.

Langkah pertama pada proses ini adalah memasang kalimat dengan tingkat kesamaan tertinggi. Pemasangan dilakukan dengan menerapkan perulangan bersarang dengan nilai batas perulangan adalah jumlah kalimat pada jawaban dan jumlah kalimat pada kunci jawaban. Dari proses ini akan didapatkan pasangan kalimat dengan kesamaan tertinggi. Setelah dilakukan perulangan, maka akan terbentuk tabel berikut (berdasarkan studi kasus)

:

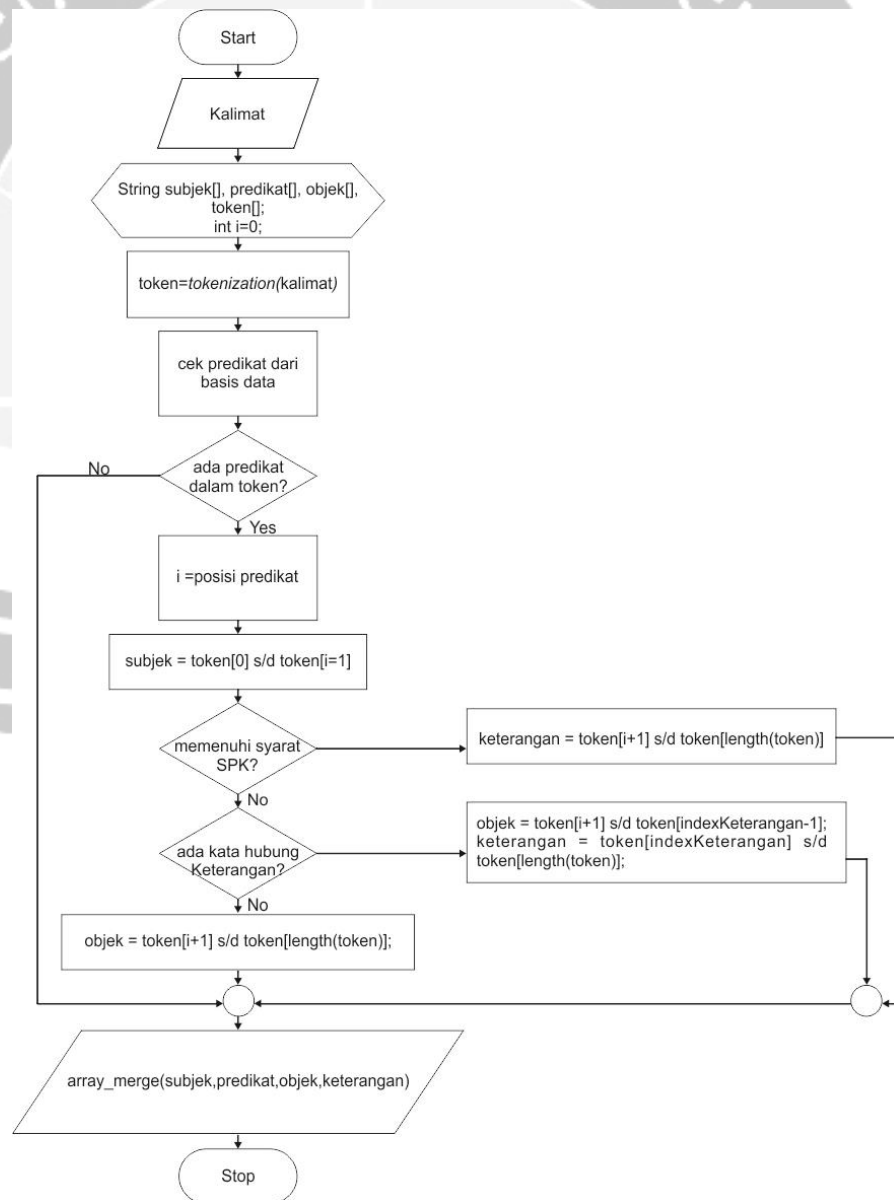
Tabel 6 Tabel perhitungan pasangan kalimat

Kunci	K1	K2	K3	K4
Jwbn 1				
K1	1	0,31	0,17	0,33
K2	0,31	1	0,22	0,44
K3	0,17	0,22	1	0,25

Kunci	K1	K2	K3	K4
Jwbn 2				
K1	0,7	0,4	0,22	0,44

Setelah didapatkan pasangan kalimat, selanjutnya akan dilakukan pengecekan struktur setiap kalimat. Struktur kalimat didapatkan dengan cara menemukan kata predikat pada kalimat. Jika ditemukan kata predikat, maka dapat ditentukan bahwa kata sebelum predikat adalah subjek. Selanjutnya akan dicek kembali apakah kalimat memenuhi syarat SPK, jika memenuhi maka kata setelah predikat akan ditentukan

sebagai keterangan. Jika kalimat tidak memenuhi syarat SPK maka akan dicek apakah ada unsur kata hubung keterangan dalam kalimat, jika ada maka kata setelah predikat akan ditentukan sebagai objek dan kata hubung sampai kata terakhir dalam kalimat ditentukan sebagai keterangan. Jika tidak ada unsur kata hubung keterangan, maka kata setelah predikat ditentukan sebagai objek. Alur proses dari penentuan struktur kalimat dapat dilihat pada gambar 6.



Gambar 6 Gambar *flowchart* proses penentuan struktur kalimat

Adapun hasil dari pengecekan struktur kalimat dari setiap kalimat adalah sebagai berikut :

Tabel 7. Tabel hasil pengecekan struktur kalimat

Kunci Jawaban	
Kalimat 1	S (Hak kepemilikan perangkat lunak) P (merupakan) K (pembatasan yang diberlakukan kepada penggunaan, penyalinan, dan modifikasi yang telah ditetapkan untuk pemegang hak dari perangkat lunak)
Kalimat 2	S (Hak kepemilikan perangkat lunak) P (menjelaskan) O (aktivitas apa saja yang diperbolehkan) K (dalam menggunakan perangkat lunak tersebut dalam bentuk lisensi)
Kalimat 3	S (Hal ini) P (menyebabkan) O (pemilik hak tidak membeli sepenuhnya)
Kalimat 4	S (Hak kepemilikan perangkat lunak) P (dibuat) O (setelah perangkat lunak atau program tersebut dikembangkan)
Jawaban 1	
Kalimat 1	S (Hak kepemilikan perangkat lunak) P (merupakan) K (pembatasan

	yang diberlakukan kepada penggunaan, penyalinan, dan modifikasi yang telah ditetapkan untuk pemegang hak dari perangkat lunak)
Kalimat 2	S (Hak kepemilikan perangkat lunak) P (menjelaskan) O (aktivitas apa saja yang diperbolehkan) K (dalam menggunakan perangkat lunak tersebut dalam bentuk lisensi)
Kalimat 3	S (Hal ini) P (menyebabkan) O (pemilik hak tidak membeli sepenuhnya)
Jawaban 2	
Kalimat 1	S (Hak kepemilikan perangkat lunak) P (merupakan) K (pembatasan penggunaan perangkat lunak oleh pemegang hak)

Langkah ketiga adalah menghitung nilai kesamaan tiap pasangan kalimat berdasarkan stukturanya. Perhitungan menggunakan perulangan yang akan mengecek setiap kata dalam struktur (S, P, O, atau K). Jika struktur dan katanya sama, maka sebuah variabel bernama count akan ditambah nilainya. Nilai kesamaan didapat dari nilai akhir count dibagi dengan gabungan dari kalimat pertama dengan kalimat kedua. Dari studi kasus, hasil dari tahapan ini akan disajikan dalam tabel berikut :

Tabel 8. Tabel hasil perhitungan kesamaan struktur kalimat

Kunci Jawaban	Jawaban 1	Sim. Struktur Kalimat
K1	K1	1
K2	K2	1
K3	K4	1
K4	K2	0,28

Kunci Jawaban	Jawaban 2	Sim. Struktur Kalimat
K1	K1	0,71
K2	K1	0,42
K3	K1	0
K4	K1	0,29

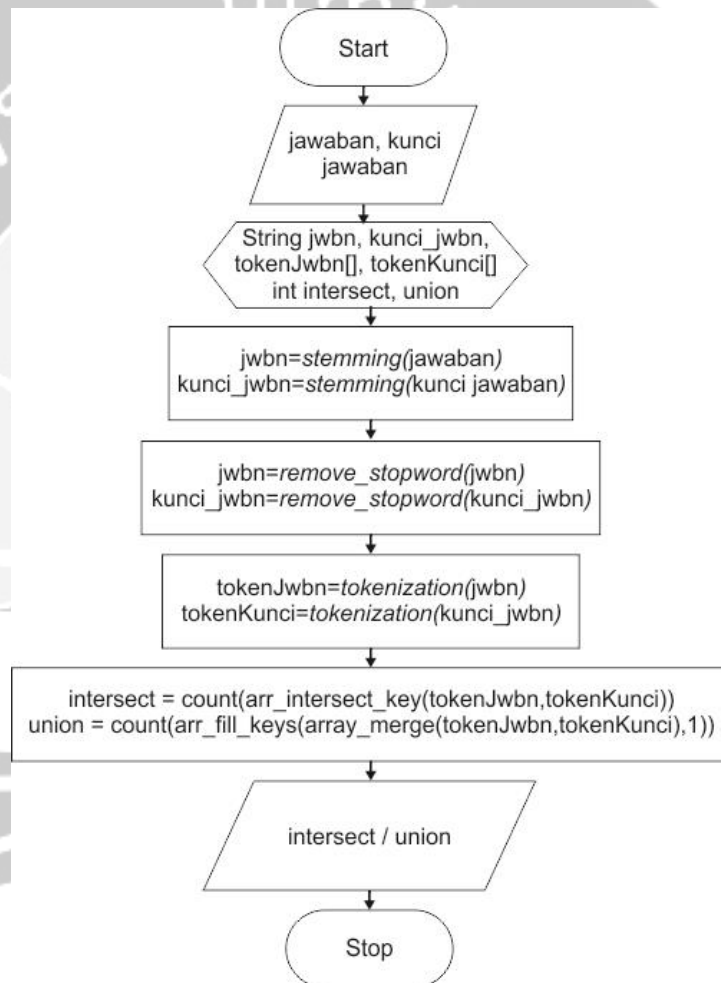
Langkah terakhir adalah menghitung rata-rata dari nilai yang telah diperoleh. Nilai rata-rata didapat dari nilai total hasil perhitungan kesamaan setiap kalimat dibagi dengan nilai terbesar dari jumlah kalimat pada kunci jawaban dan jumlah kalimat pada jawaban. Pada studi kasus, nilai terbesar dari jumlah kalimat adalah jumlah kalimat kunci jawaban yaitu 4 sehingga didapatkan :

$$\text{Sim}(\text{Kunci}, \text{Jwb1}) = \frac{1+1+1+0,28}{4} = 0,82$$

$$\text{Sim}(\text{Kunci}, \text{Jwb2}) = \frac{0,71+0,42+0+0,29}{4} = 0,355$$

Jadi pada proses ini, nilai yang diperoleh untuk jawaban 1 adalah 0,82 dan nilai yang diperoleh jawaban 2 adalah 0,355.

3.2. Proses Perhitungan Nilai Kesamaan Kalimat dengan Algoritma Jaccard



Gambar 7 Gambar *flowchart* perhitungan kesamaan kalimat

Pada proses ini, pertama akan dilakukan proses *stemming* untuk jawaban dan kunci jawaban. Dari studi kasus hasil untuk proses *stemming* kunci

jawaban, jawaban 1 dan jawaban 2 adalah sebagai berikut

Tabel 9. Tabel hasil *stemming* kalimat kunci jawaban dan jawaban

Kunci Jawaban	hak milik perangkat lunak rupa batas yang laku kepada guna salin dan modifikasi yang telah tetap untuk pegang hak dari perangkat lunak. hak milik perangkat lunak jelas aktivitas apa saja yang boleh dalam guna perangkat lunak sebut dalam bentuk lisensi. hal ini sebab milik hak tidak beli sepuh. hak milik perangkat lunak buat telah perangkat lunak atau program sebut kembang.
Jawaban 1	hak milik perangkat lunak rupa batas yang laku kepada guna salin dan modifikasi yang telah tetap untuk pegang hak dari perangkat lunak. hak milik perangkat lunak jelas aktivitas apa saja yang boleh dalam guna perangkat lunak sebut dalam bentuk lisensi. hal ini sebab milik hak tidak beli sepuh
Jawaban 2	hak milik perangkat lunak rupa batas guna perangkat lunak oleh pegang hak

Proses selanjutnya adalah menghilangkan *stopwords* atau kata henti yang sudah disimpan dalam basis data. Setelah menghilangkan *stopwords* kunci

jawaban, jawaban 1, dan jawaban 2 akan menjadi seperti berikut :

Tabel 10 Tabel hasil pengapusan *stopwords* kalimat kunci jawaban dan jawaban

Kunci Jawaban	hak milik perangkat lunak rupa batas laku salin modifikasi pegang hak perangkat lunak hak milik perangkat lunak aktivitas perangkat lunak bentuk lisensi milik hak beli sepuh hak milik perangkat lunak perangkat lunak program kembang
Jawaban 1	hak milik perangkat lunak rupa batas laku salin modifikasi pegang hak perangkat lunak hak milik perangkat lunak aktivitas perangkat lunak bentuk lisensi milik hak beli sepuh
Jawaban 2	hak milik perangkat lunak rupa batas perangkat lunak pegang hak

Langkah ketiga adalah melakukan proses *tokenization*. Proses ini akan memecah kalimat jawaban dan kunci jawaban menjadi kata. Setelah proses *tokenization*, akan ditentukan irisan dan gabungan dari setiap kata pada jawaban dan kunci jawaban, berikut hasil dari proses ini

Tabel 11. Tabel irisan dan gabungan kunci jawaban dengan jawaban

Kunci Jawaban dengan Jawaban 1		Jumlah
Irisan	hak, milik, perangkat, lunak, rupa, batas, laku, salin, modifikasi, pegang, aktifitas, bentuk, lisensi, beli, sepuh.	15
Gabungan	hak, milik, perangkat, lunak, rupa, batas, laku, salin, modifikasi, pegang, aktifitas, bentuk, lisensi, beli, sepuh, program, kembang.	17
Kunci Jawaban dengan Jawaban 2		
Irisan	Hak, milik, perangkat, lunak, rupa, batas, pegang.	7
Gabungan	hak, milik, perangkat, lunak, rupa, batas, laku, salin, modifikasi, pegang, aktifitas, bentuk, lisensi, beli, sepuh, program, kembang.	17

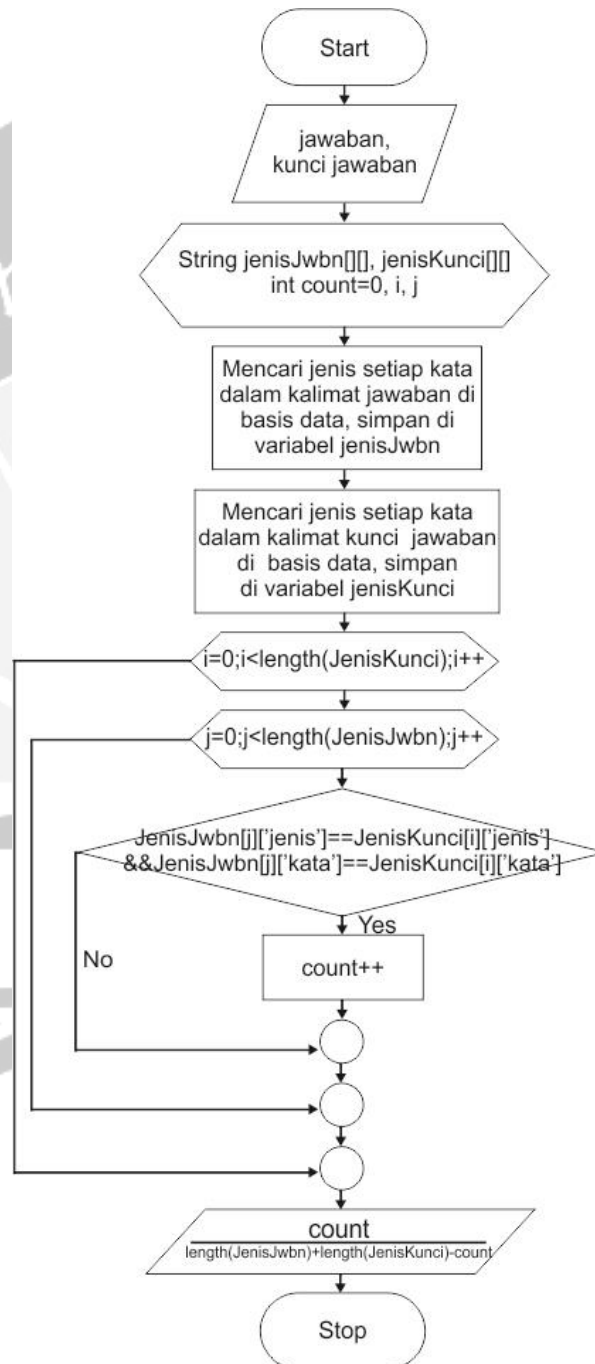
Selanjutnya hitung nilai jaccard dari tabel hasil diatas. Rumus jaccard adalah irisan dibagi dengan gabungan, sehingga didapatkan :

$$J(\text{kunci, jwb1}) = \frac{15}{17} = 0,88$$

$$J(\text{kunci, jwb2}) = \frac{7}{17} = 0,41$$

Jadi pada proses ini, nilai yang diperoleh untuk jawaban 1 adalah 0,88 dan nilai yang diperoleh jawabah 2 adalah 0,41.

3.3. Proses Perhitungan Nilai Kesamaan Jenis Kata



Gambar 8 Gambar *flowchart* perhitungan kesamaan jenis kata

Masukan dari proses ini adalah jawaban dan kunci jawaban yang sudah melalui proses *tokenizer* dan

penghapusan *stopwords*. Proses ini dimulai dengan mencari jenis dari setiap kata pada kalimat jawaban dan kunci jawaban di basis data. Dari studi kasus yang dibuat, didapatkan jenis setiap kata pada kalimat adalah sebagai berikut :

Tabel 11. Tabel jenis kalimat jawaban dan kunci jawaban

Kunci jawaban	Hak=>/N, perangkat=>/N, pembatasan=>/N, penggunaan=>/N, modifikasi=>v, pemegang=>/N, diperbolehkan=>/N, lisensi=>n, pemilik=>/N, sepenuhnya=>n, dikembangkan=>n	kepemilikan=>n, lunak=>n, diberlakukan=>n, penyalinan=>/N, ditetapkan=>/N, aktivitas=>n, bentuk=>adj, menyebabkan=>n, membeli=>/N, program=>/N,
Jawaban 1	Hak=>/N, perangkat=>/N, pembatasan=>/N, penggunaan=>/N, modifikasi=>v, pemegang=>/N, diperbolehkan=>/N, lisensi=>n, pemilik=>/N, sepenuhnya=>n	kepemilikan=>n, lunak=>n, diberlakukan=>n, penyalinan=>/N, ditetapkan=>/N, aktivitas=>n, bentuk=>adj, menyebabkan=>n, membeli=>/N,
Jawaban 2	Hak=>/N, perangkat=>/N, pembatasan=>/N, pemegang=>/N, hak=>/N	kepemilikan=>n, lunak=>n, penggunaan=>/N,

Langkah selanjutnya adalah mencari irisan kata dan gabungan kata dengan jenis yang sama dari jawaban dan kunci jawaban. Pada implementasi program, penulis menggunakan perulangan bersarang untuk menentukan jumlah irisan dan kemudian menggunakan jumlah irisan tersebut untuk menentukan jumlah gabungan kata. Adapun hasil dari proses ini berdasarkan studi kasus adalah sebagai berikut :

Tabel 12 Tabel hasil irisan dan gabungan jenis kata

	Kunci Jawaban dengan Jawaban 1	Jumlah
Irisan	Hak=>/N, kepemilikan=>n, perangkat=>/N, lunak=>n, pembatasan=>/N, diberlakukan=>n, penggunaan=>/N, penyalinan=>/N, modifikasi=>v, ditetapkan=>/N, pemegang=>/N, hak=>/N, aktivitas=>n, diperbolehkan=>/N, bentuk=>adj, lisensi=>n, Hal=>/N, menyebabkan=>n, pemilik=>/N, membeli=>/N, sepenuhnya=>n	21
Gabungan	Hak=>/N, kepemilikan=>n, perangkat=>/N, lunak=>n, pembatasan=>/N, diberlakukan=>n, penggunaan=>/N, penyalinan=>/N, modifikasi=>v, ditetapkan=>/N, pemegang=>/N, hak=>/N, aktivitas=>n, diperbolehkan=>/N, bentuk=>adj, lisensi=>n, Hal=>/N, menyebabkan=>n, pemilik=>/N,	23

	membeli=>/N, sepenuhnya=>n, program=>/N, dikembangkan=>n	
Kunci Jawaban dengan Jawaban 2		
Irisan	Hak=>/N, kepemilikan=>n, perangkat=>/N, lunak=>n, pembatasan=>/N, penggunaan=>/N, pemegang=>/N, hak=>/N	8
Gabungan	Hak=>/N, kepemilikan=>n, perangkat=>/N, lunak=>n, pembatasan=>/N, diberlakukan=>n, penggunaan=>/N, penyalinan=>/N, modifikasi=>v, ditetapkan=>/N, pemegang=>/N, hak=>/N, aktivitas=>n, diperbolehkan=>/N, bentuk=>adj, lisensi=>n, Hal=>/N, menyebabkan=>n, pemilik=>/N, membeli=>/N, sepenuhnya=>n, program=>/N, dikembangkan=>n	23

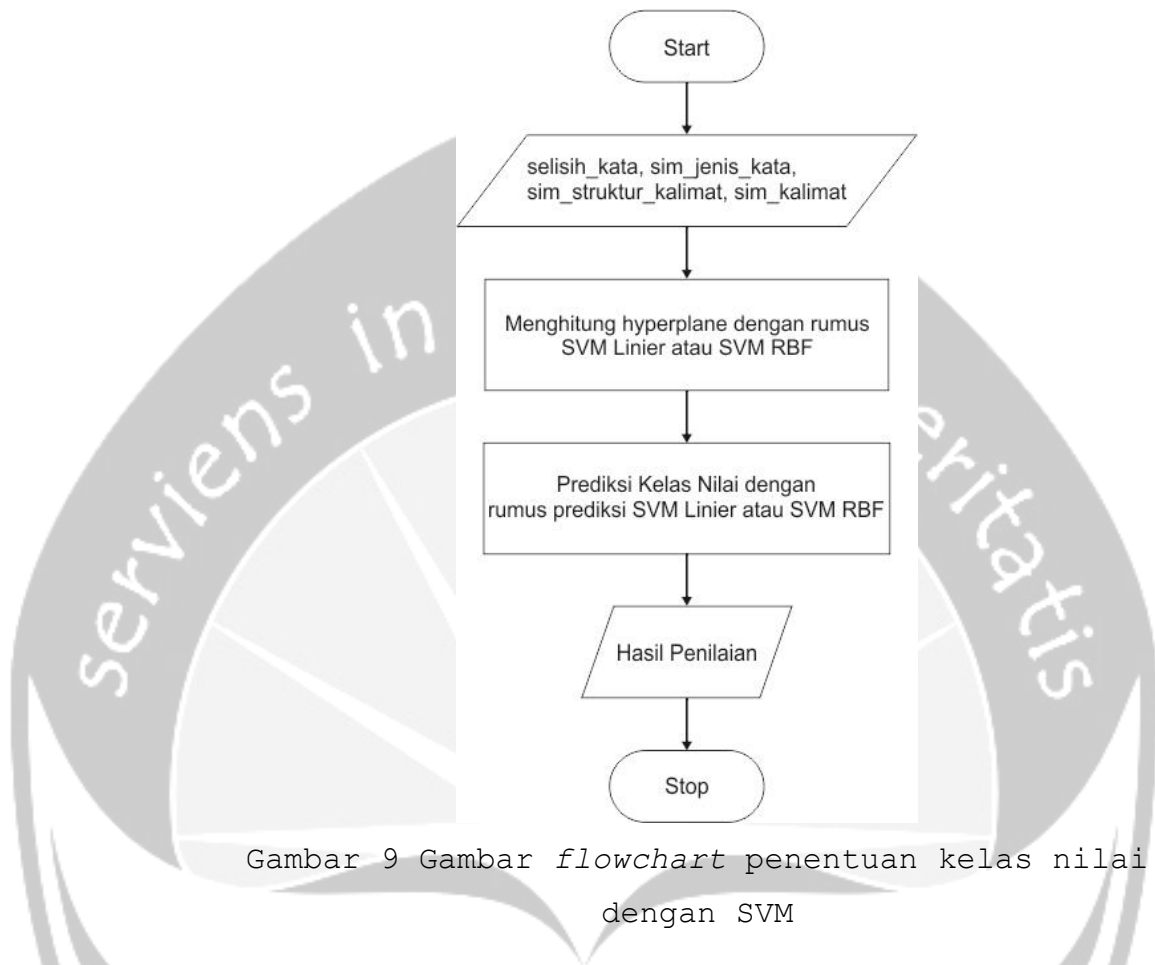
Hasil dari perhitungan adalah sebagai berikut:

$$J(\text{kunci}, \text{jwb1}) = \frac{21}{23} = 0,91$$

$$J(\text{kunci}, \text{jwb2}) = \frac{8}{23} = 0,34$$

Jadi pada proses ini, nilai yang diperoleh untuk jawaban 1 adalah 0,91 dan nilai yang diperoleh jawaban 2 adalah 0,34.

3.4. Proses Penentuan Kelas Nilai dengan SVM



Gambar 9 Gambar *flowchart* penentuan kelas nilai dengan SVM

Masukan dari proses ini adalah hasil dari proses penghitungan selisih kata, penghitungan nilai kesamaan jenis kata, penghitungan kesamaan jenis kalimat dan penghitungan kesamaan kalimat. Misalkan kita memiliki nilai input untuk jawaban 1 adalah :

Kesamaan jenis kata (x1) = 0,91
Kesamaan kalimat (x2) = 0,88
Kesamaan jenis kalimat (x3) = 0,82
Selisih jumlah kata (x4) = 0,02

Misalkan kita akan menghitung kelas yang dimiliki oleh jawaban tersebut menggunakan SVM linier dengan jumlah kelas adalah 2 (kelas A (+1)

dan kelas B (-1)) dan diketahui hasil dari perhitungan hyperplane dengan SVM linier mendapatkan nilai $w = 1$ dan $b = -2,2$, sehingga terbentuk persamaan :

$$f(x) = x_1 + x_2 + x_3 + x_4 - 2,2 = 0$$

Maka dapat dihitung nilai peramalan data dengan rumus $\text{sign}(f(x))$.

$$\text{Hasil} = \text{sign}(x_1 + x_2 + x_3 + x_4 - 2,2)$$

$$\text{Hasil} = \text{sign}(0,91 + 0,88 + 0,82 + 0,02 - 2,2)$$

$$\text{Hasil} = \text{sign}(0,43)$$

$$\text{Hasil} = +1$$

Jadi kelas yang dimiliki oleh jawaban 1 adalah +1 yang merupakan kelas A. Kemudian dengan kondisi yang sama diketahui bahwa nilai dari kelas 2 adalah :

$$\text{Kesamaan jenis kata } (x_1) = 0,34$$

$$\text{Kesamaan kalimat } (x_2) = 0,41$$

$$\text{Kesamaan jenis kalimat } (x_3) = 0,355$$

$$\text{Selisih jumlah kata } (x_4) = 0,9$$

Maka hasil dari peramalan dengan SVM linier adalah :

$$\text{Hasil} = \text{sign}(x_1 + x_2 + x_3 + x_4 - 2,2)$$

$$\text{Hasil} = \text{sign}(0,34 + 0,41 + 0,355 + 0,9 - 2,2)$$

$$\text{Hasil} = \text{sign}(-0,195)$$

$$\text{Hasil} = -1$$

Jadi kelas yang dimiliki oleh jawaban 2 adalah -1 yang merupakan kelas B.