

BAB I

PENDAHULUAN

1.1 Latar Belakang

Al-Qur'an adalah buku klasik dan bahasanya merupakan bahasa Arab tradisional yang dikenal sebagai i'rab. Al-Qur'an memiliki kelebihan, yaitu merupakan bentuk korpus tertutup, yang meliputi: Pertama, memiliki pengulangan struktur yang sering dari frasa yang sama. Kedua, Al-Qur'an secara tradisional diidentifikasi dengan satu orang, wilayah tertentu, dan jangka waktu tertentu dan jumlahnya relatif terbatas. Kedua fakta ini membenarkan perlakuan terhadap Al-Qur'an sebagai korpus independen yang layak mendapatkan studi independen bahasanya secara umum dan sintaksis pada khususnya. Memahami Al-Qur'an adalah tantangan besar bagi masyarakat, untuk pendidikan umum barat, pendidikan Muslim di dunia, representasi pengetahuan dan penalaran, pengetahuan ekstraksi dari teks, kebenaran, dan kolaborasi online. Memahami Al-Qur'an adalah hal yang utama dan merupakan tantangan besar untuk ilmu komputer dan kecerdasan buatan [1].

Dalam era digital saat ini, akses terhadap kitab suci Al-Qur'an semakin mudah melalui platform aplikasi, situs web, maupun media digital lainnya. Namun demikian, pemanfaatan ini belum sepenuhnya menyentuh nilai emosional dan psikologis pengguna, khususnya dalam hal menemukan ayat yang sesuai dengan suasana hati mereka. Ketika seseorang berada dalam kondisi sedih, marah, atau cemas, menemukan ayat yang tepat secara manual seringkali tidak efisien dan kurang personal. Oleh karena itu, dibutuhkan sistem yang dapat secara otomatis

menganalisis suasana hati pengguna melalui teks dan merekomendasikan ayat Al-Qur'an yang relevan secara tepat dan cepat. Pendekatan ini memiliki potensi tinggi dalam memberikan dukungan spiritual berbasis konteks emosional pengguna.

Analisis suasana hati atau emosi pengguna dapat dilakukan dengan mengandalkan algoritma klasifikasi teks berbasis *NLP*. Salah satu algoritma yang terbukti efektif adalah *Support Vector Machine (SVM)*. *SVM* mampu menangani data teks berdimensi tinggi dan bekerja dengan membangun hyperplane yang memisahkan kategori-kategori emosi, seperti senang, sedih, marah, takut, atau tenang. Penelitian sebelumnya menunjukkan bahwa *SVM* memiliki performa yang kompetitif, bahkan pada dataset yang relatif kecil, dan seringkali lebih stabil dibanding algoritma lain seperti *Naïve Bayes* atau *K-Nearest Neighbors*. Hal ini menjadikan *SVM* sebagai pilihan tepat untuk mengklasifikasikan teks masukan pengguna sebelum diterjemahkan menjadi rekomendasi ayat Al-Qur'an [2].

Selain keunggulan teknis, penerapan sistem rekomendasi ayat Al-Qur'an berbasis analisis suasana hati juga menawarkan manfaat praktis bagi pengguna. Sistem ini dapat membantu umat Islam menemukan ayat yang sesuai dengan keadaan emosional mereka, sehingga mendukung refleksi diri, ketenangan hati, dan penguatan spiritual. Misalnya, pengguna yang sedang merasa sedih dapat secara otomatis mendapatkan ayat-ayat yang memberikan ketenangan dan motivasi, sementara pengguna yang merasa bahagia dapat diarahkan kepada ayat yang mengajak bersyukur. Dengan demikian, sistem ini tidak hanya berperan sebagai

sarana informasi, tetapi juga sebagai media bimbingan spiritual yang personal dan relevan dengan kondisi pengguna

Analisis teks emosional seperti klasifikasi sentiment/emosi merupakan bagian penting dalam *Natural Language Processing (NLP)*. Algoritma klasik seperti *Support Vector Machine (SVM)* terbukti efektif dalam menangani dimensi tinggi seperti *TF-IDF*, serta mendapatkan performa yang kompetitif di berbagai kasus teks. Sebagai contoh, *SVM* berhasil digunakan untuk analisis sentimen status Twitter Bahasa Indonesia, memperoleh akurasi sekitar 77 % dan F-measure sekitar 68 % setelah ditambah langkah *Query Expansion Ranking* [3]. Studi lainnya dari *Journal of Advances in Information Technology* (2024) membandingkan berbagai model klasifikasi emosi pada teks Bahasa Indonesia dengan *SVM*. Hasil menunjukkan bahwa variasi kernel seperti *Polynomial* atau *RBF* memberikan peningkatan dalam klasifikasi emosi multikategori dibanding *SVM linear* [4].

SVM juga sering ditemui sebagai model paling populer dalam literatur klasifikasi emosi berbasis teks. Dalam sebuah tinjauan sistematis terhadap algoritma deteksi emosi teks, ditemukan bahwa 78 % penelitian menggunakan *SVM*, menjadikannya algoritma *top-performing* dalam domain tersebut [5].

Lebih jauh, implementasi *SVM* juga efektif dalam skenario *real-world*. Misalnya, studi mengenai sentimen pengguna Twitter terhadap pergantian pelatih tim nasional Indonesia menggunakan *TF-IDF* dan *SVM* menghasilkan akurasi sebesar 83 %, with precision 82 % dan recall 84 % [6]. Ini memperlihatkan bahwa *SVM* mampu bekerja secara konsisten dengan hasil evaluasi yang baik dalam data teks sosial media yang nyata digunakan oleh masyarakat Indonesia.

Penelitian yang dilakukan oleh Moch. Fauzan dkk (2022) dengan judul Klasifikasi Al-Qur'an Terjemahan Bahasa Indonesia Dengan Menggunakan Algoritma *Support Vector Machine (SVM)* menyatakan bahwa *SVM* tanpa *feature selection* memberi performa terbaik dengan AUC 83,3%. Artinya *SVM* cukup efektif mengklasifikasikan ayat ke dalam kategori topik tertentu. Termasuk juga penelitian oleh Annisa dkk (2022) dengan judul identifikasi ayat sulit dalam al-quran menggunakan *SVM*, *LSTM*, dan *CNN* menyatakan bahwa *SVM* menghasilkan akurasi tertinggi (68%), lebih baik daripada *CNN* (66%) dan *LSTM* (65%), sehingga *SVM* dinilai cocok untuk klasifikasi ayat sulit vs mudah. Kemudian penelitian yang dilakukan oleh Novi Amalia Hidayat (2020) dengan judul klasifikasi teks hasil tafsir ayat sains dalam al-qur'an menggunakan algoritma *SVM* menyimpulkan bahwasanya *SVM* mampu mengidentifikasi ayat tafsir bertema sains secara lebih terstruktur sehingga mempermudah pencarian ayat terkait ilmu pengetahuan.

Dengan latar belakang tersebut, penelitian ini bertujuan untuk mengembangkan sistem rekomendasi ayat Al-Qur'an otomatis yang didukung analisis suasana hati pengguna menggunakan NLP dan algoritma *SVM*. Sistem ini diharapkan menjadi inovasi di bidang teknologi informasi dan keagamaan, yang tidak hanya mempermudah akses terhadap ayat Al-Qur'an, tetapi juga memberikan manfaat psikologis dan spiritual kepada masyarakat. Implementasi sistem ini dapat menjadi solusi praktis bagi umat Islam di era digital untuk mendapatkan bimbingan spiritual secara tepat, cepat, dan sesuai kondisi emosional mereka.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, rumusan masalah dalam penelitian ini adalah Bagaimana merancang dan mengimplementasikan sistem rekomendasi ayat Al-Qur'an otomatis berdasarkan analisis suasana hati pengguna menggunakan *Natural Language Processing (NLP)* dan algoritma klasifikasi *SVM*?

1.3 Tujuan Penelitian

Tujuan dari penelitian ini adalah untuk merancang dan mengimplementasikan sistem rekomendasi ayat Al-Qur'an otomatis berdasarkan analisis suasana hati pengguna menggunakan *Natural Language Processing (NLP)* dan algoritma klasifikasi *SVM*.

1.4 Batasan Masalah

Agar pembahasan masalah menjadi lebih jelas dan terarah, maka diperlukan adanya batasan masalah. Adapun Batasan penelitian ini adalah:

1. Analisis hanya dilakukan pada teks berbahasa Indonesia. Teks dalam bahasa lain tidak diproses dalam penelitian ini.
2. Sistem hanya difokuskan pada pendeteksian emosi manusia berdasarkan teks yang diberikan pengguna, dan tidak digunakan untuk mendeteksi emosi makhluk hidup lain seperti hewan atau entitas non-manusia.
3. Suasana hati dikategorikan menjadi beberapa label tertentu yaitu cinta, marah, sedih, senang, dan takut.
4. Algoritma yang digunakan untuk proses klasifikasi adalah *Support Vector Machine (SVM)*, tanpa membandingkan performanya dengan algoritma lain.

5. Sistem hanya menampilkan rekomendasi ayat Al-Qur'an yang relevan dengan hasil klasifikasi suasana hati pengguna berdasarkan kesesuaian tema, tanpa memberikan tafsir mendalam atau penjelasan hukum fiqih terkait ayat tersebut.
6. Implementasi sistem dilakukan berbasis aplikasi web menggunakan bahasa pemrograman *Python* dan *framework streamlit*

1.5 Manfaat penelitian

Penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1. Sistem yang dibangun dapat digunakan oleh masyarakat untuk memperoleh ayat-ayat Al-Qur'an yang relevan dengan suasana hati mereka, sehingga dapat memberikan ketenangan batin, motivasi, dan bimbingan moral secara cepat dan tepat.
2. dapat menjadi referensi bagi peneliti lain yang ingin mengembangkan sistem serupa dengan algoritma atau metode yang berbeda, ataupun memperluas cakupan analisis ke bahasa lain dan sumber teks religius lainnya.

1.6 Sistematika Penulisan

Sistematika penulisan dalam skripsi ini disusun untuk memudahkan pembaca memahami alur penelitian dari awal hingga akhir. Adapun sistematika penulisan yang digunakan adalah sebagai berikut:

BAB I PENDAHULUAN

Bab ini berisi uraian mengenai latar belakang masalah, rumusan masalah, tujuan penelitian, batasan masalah, manfaat penelitian, dan sistematika

penulisan. Bab ini memberikan gambaran umum mengenai alasan dilakukannya penelitian serta arah yang akan dicapai.

BAB II LANDASAN TEORI

Bab ini mengulas berbagai teori dan konsep yang relevan, seperti *Natural Language Processing (NLP)*, analisis sentimen, *Support Vector Machine (SVM)*, dan konsep ayat Al-Qur'an.

BAB III METODOLOGI PENELITIAN

Bab ini menjelaskan metode penelitian yang digunakan, meliputi pengamatan terdahulu sampai kesimpulan.

BAB IV ANALISIS DAN PERANCANGAN SISTEM

Bab ini memaparkan hasil implementasi sistem rekomendasi ayat Al-Qur'an berbasis analisis suasana hati menggunakan *NLP* dan *SVM*. Bagian ini dimulai dengan penjelasan mengenai proses pengolahan data teks, seperti *preprocessing* (tokenisasi, *stemming*, *stopword removal*), ekstraksi fitur menggunakan metode yang sesuai, dan pelatihan model menggunakan algoritma *SVM*. Selanjutnya disajikan hasil pengujian sistem, baik dari sisi akurasi algoritma maupun kualitas rekomendasi ayat yang diberikan. Pembahasan meliputi analisis hasil pengujian yang dibandingkan dengan penelitian sebelumnya, kelebihan dan kekurangan sistem yang dibangun, serta evaluasi performa sistem berdasarkan parameter yang telah ditentukan.

BAB V IMPLEMENTASI DAN PENGUJIAN

Implementasi sistem merupakan tahap realisasi dari hasil perancangan yang telah dibuat pada bab sebelumnya. Sistem rekomendasi ayat Al-Qur'an ini

diimplementasikan menggunakan bahasa pemrograman *Python* dengan pustaka pendukung seperti *NLTK* untuk pemrosesan bahasa alami (*Natural Language Processing*), *scikit-learn* untuk penerapan algoritma *Support Vector Machine (SVM)*.

BAB VI KESIMPULAN DAN SARAN

Bab ini berisi kesimpulan yang merupakan jawaban atas rumusan masalah penelitian, yang diperoleh dari hasil analisis dan pengujian sistem. Selain itu, diberikan pula saran untuk pengembangan penelitian di masa yang akan datang agar sistem yang dibuat dapat lebih optimal, baik dari segi akurasi, kelengkapan dataset, maupun kenyamanan pengguna.

BAB II

LANDASAN TEORI

2.1 Alqur'an

2.1.1 Definisi Etimologis dan Terminologis

Secara etimologis, kata *Al-Qur'an* berasal dari bahasa Arab qara'a–yaqra'u–qur'an yang berarti membaca atau bacaan. Dengan demikian, Al-Qur'an bermakna “sesuatu yang dibaca berulang-ulang”. Hal ini sesuai dengan sifat Al-Qur'an yang diturunkan untuk dibaca, dihafal, dan diamalkan [7].

Secara terminologis, Al-Qur'an didefinisikan sebagai kalam Allah SWT yang diturunkan kepada Nabi Muhammad SAW melalui malaikat Jibril, ditulis dalam mushaf, diriwayatkan secara mutawatir, serta membacanya bernilai ibadah. Definisi ini menegaskan otoritas dan keaslian Al-Qur'an sebagai wahyu ilahi yang berlaku sepanjang masa [8].

Secara etimologis dan terminologis, Al-Qur'an tidak hanya dipahami sebagai bacaan yang dilafalkan, tetapi juga sebagai wahyu ilahi yang mengandung petunjuk hidup bagi manusia. Makna etimologis Al-Qur'an menekankan fungsinya sebagai bacaan yang dibaca, dihafal, dan direnungkan secara berulang, sedangkan secara terminologis Al-Qur'an dipahami sebagai kalam Allah SWT yang diturunkan kepada Nabi Muhammad SAW melalui malaikat Jibril, diriwayatkan secara mutawatir, ditulis dalam mushaf, serta membacanya bernilai ibadah. Pemahaman ini menegaskan bahwa Al-Qur'an memiliki kedudukan sebagai sumber utama ajaran Islam yang bersifat transenden sekaligus kontekstual, sehingga relevan

dijadikan landasan nilai dalam berbagai bidang kehidupan, termasuk kajian akademik dan pengembangan ilmu pengetahuan modern [9].

2.1.2 Kedudukan Al-Qur'an

Al-Qur'an merupakan kitab suci terakhir yang menjadi pedoman hidup umat Islam. Kedudukannya mencakup berbagai aspek, antara lain: (1) sebagai petunjuk hidup (*hudan*), (2) sebagai pembeda (*furqan*) antara yang benar dan salah, (3) sebagai rahmat dan penawar (*syifa' wa rahmah*) bagi orang beriman, dan (4) sebagai sumber hukum Islam. Hal ini menunjukkan bahwa Al-Qur'an bukan sekadar teks ritual, melainkan fondasi dalam seluruh aspek kehidupan [7].

Al-Qur'an memiliki kedudukan utama sebagai sumber hukum Islam yang pertama dan paling otoritatif. Seluruh aspek ajaran Islam, baik yang berkaitan dengan akidah, ibadah, muamalah, maupun akhlak, bersumber dari Al-Qur'an sebagai wahyu Allah SWT. Kedudukannya tidak dapat digantikan oleh sumber hukum lain, karena Al-Qur'an menjadi dasar bagi hadis, *ijma'*, dan *qiyas*. Oleh karena itu, setiap pengembangan hukum, pemikiran, dan praktik keislaman harus merujuk dan tidak boleh bertentangan dengan prinsip-prinsip yang terkandung dalam Al-Qur'an [10].

Selain sebagai sumber hukum, Al-Qur'an juga berkedudukan sebagai pedoman hidup (*way of life*) bagi umat manusia yang bersifat universal dan sepanjang masa. Kandungan Al-Qur'an tidak hanya mengatur hubungan manusia dengan Allah SWT (*hablum minallah*), tetapi juga hubungan antar sesama manusia (*hablum minannas*) serta hubungan manusia dengan lingkungan. Kedudukan ini menjadikan Al-Qur'an relevan dalam berbagai konteks kehidupan modern,

termasuk dalam bidang pendidikan, sosial, psikologi, dan teknologi informasi. Dengan demikian, Al-Qur'an dapat dijadikan dasar nilai dan etika dalam pengembangan sistem berbasis teknologi, seperti sistem rekomendasi yang berorientasi pada kebutuhan spiritual dan emosional manusia [11].

2.1.3 Karakteristik Al-Qur'an

Al-Qur'an memiliki beberapa karakteristik utama, di antaranya: (1) universal, berlaku untuk semua manusia tanpa batasan ruang dan waktu; (2) otentik, terjaga keasliannya sejak diturunkan hingga akhir zaman (QS. Al-Hijr [15]:9); (3) mukjizat, tidak dapat ditandingi dari segi bahasa, sastra, dan kandungan makna; serta (4) komprehensif, karena mencakup ajaran akidah, ibadah, muamalah, dan akhlak. Karakteristik ini menjadikan Al-Qur'an relevan untuk setiap kondisi kehidupan manusia [12].

Salah satu karakteristik utama Al-Qur'an adalah sifatnya yang bersifat universal dan kontekstual. Universal berarti ajaran Al-Qur'an ditujukan untuk seluruh umat manusia tanpa dibatasi oleh ruang, waktu, ras, maupun budaya tertentu. Di sisi lain, kontekstual menunjukkan bahwa kandungan Al-Qur'an dapat dipahami dan diterapkan sesuai dengan perkembangan zaman dan kondisi sosial masyarakat. Fleksibilitas makna ini menjadikan Al-Qur'an tetap relevan dalam menghadapi berbagai persoalan modern, baik dalam bidang sosial, pendidikan, ekonomi, maupun teknologi, tanpa kehilangan nilai dan prinsip dasarnya [13].

Karakteristik penting lainnya dari Al-Qur'an adalah kesempurnaan dan keseimbangan kandungannya. Al-Qur'an memuat ajaran yang menyeluruh, mencakup aspek spiritual, moral, hukum, dan kemanusiaan secara seimbang. Selain

itu, Al-Qur'an juga memiliki konsistensi isi dan keindahan bahasa yang tinggi, sehingga mampu menyentuh aspek intelektual dan emosional pembacanya. Keunikan struktur bahasa serta kedalaman makna Al-Qur'an menjadikannya tidak hanya sebagai sumber ajaran agama, tetapi juga sebagai objek kajian ilmiah lintas disiplin, termasuk linguistik, psikologi, dan ilmu komputer [14].

2.1.4 Ayat Al-Qur'an Terkait Suasana Hati

1. Sedih

QS. Al-Baqarah [2]: 286

لَا يُكَلِّفُ اللَّهُ نَفْسًا إِلَّا وُسْعَهَا ۚ لَهَا مَا كَسَبَتْ وَعَلَيْهَا مَا اكْتَسَبَتْ ۗ رَبَّنَا لَا تُؤَاخِذْنَا إِنْ نَسِينَا أَوْ أَخْطَأْنَا ۗ رَبَّنَا وَلَا تَحْمِلْ عَلَيْنَا إَصْرًا كَمَا حَمَلْتَهُ عَلَى الَّذِينَ مِنْ قَبْلِنَا ۗ رَبَّنَا وَلَا تُحَمِّلْنَا مَا لَا طَاقَةَ لَنَا بِهِ ۚ وَاعْفُ عَنَّا وَاعْفِرْ لَنَا ۗ وَارْحَمْنَا ۚ أَنْتَ مَوْلَانَا فَانصُرْنَا عَلَى الْقَوْمِ الْكَافِرِينَ ۝

286. Allah tidak membebani seseorang, kecuali menurut kesanggupannya. Baginya ada sesuatu (pahala) dari (kebajikan) yang diusahakannya dan terhadapnya ada (pula) sesuatu (siksa) atas (kejahatan) yang diperbuatnya. (Mereka berdoa,) “Wahai Tuhan kami, janganlah Engkau hukum kami jika kami lupa atau kami salah. Wahai Tuhan kami, janganlah Engkau bebani kami dengan beban yang berat sebagaimana Engkau bebankan kepada orang-orang sebelum kami. Wahai Tuhan kami, janganlah Engkau pikulkan kepada kami apa yang tidak sanggup kami memikulnya. Maafkanlah kami, ampunilah kami, dan rahmatilah kami. Engkaulah pelindung kami. Maka, tolonglah kami dalam menghadapi kaum kafir.”

2. Marah

QS. Ali Imran [3]: 134

الَّذِينَ يُنْفِقُونَ فِي السَّرَّاءِ وَالضَّرَّاءِ وَالْكُظُمِينَ الْغَيْظِ وَالْعَافِينَ عَنِ النَّاسِ ۗ وَاللَّهُ يُحِبُّ الْمُحْسِنِينَ

134. (yaitu) orang-orang yang selalu berinfaq, baik di waktu lapang maupun sempit, orang-orang yang mengendalikan kemurkaannya, dan orang-orang yang

memaafkan (kesalahan) orang lain. Allah mencintai orang-orang yang berbuat kebaikan.

3. Takut

QS. Al-Baqarah [2]: 155

وَلَنَبْلُوَنَّكُمْ بِشَيْءٍ مِّنَ الْخَوْفِ وَالْجُوعِ وَنَقْصٍ مِّنَ الْأَمْوَالِ وَالْأَنْفُسِ وَالثَّمَرَاتِ وَبَشِّرِ الصَّابِرِينَ

155. Kami pasti akan mengujimu dengan sedikit ketakutan dan kelaparan, kekurangan harta, jiwa, dan buah-buahan. Sampaikanlah (wahai Nabi Muhammad,) kabar gembira kepada orang-orang sabar,

4. Senang

QS. Yunus [10]: 58

قُلْ بِفَضْلِ اللَّهِ وَبِرَحْمَتِهِ فَبِذَلِكَ فَلْيَفْرَحُوا هُوَ خَيْرٌ مِّمَّا يَجْمَعُونَ

58. Katakanlah (Nabi Muhammad), “Dengan karunia Allah dan rahmat-Nya itu, hendaklah mereka bergembira. Itu lebih baik daripada apa yang mereka kumpulkan.”

5. Cinta

QS. Ar-Rum [30]: 21

وَمِنْ آيَاتِهِ أَنْ خَلَقَ لَكُمْ مِنْ أَنْفُسِكُمْ أَزْوَاجًا لِتَسْكُنُوا إِلَيْهَا وَجَعَلَ بَيْنَكُمْ مَوَدَّةً وَرَحْمَةً إِنَّ فِي ذَلِكَ لَآيَاتٍ لِّقَوْمٍ يَتَفَكَّرُونَ

21. Di antara tanda-tanda (kebesaran)-Nya ialah bahwa Dia menciptakan pasangan-pasangan untukmu dari (jenis) dirimu sendiri agar kamu merasa tenteram kepadanya. Dia menjadikan di antaramu rasa cinta dan kasih sayang. Sesungguhnya pada yang demikian itu benar-benar terdapat tanda-tanda (kebesaran Allah) bagi kaum yang berpikir.

2.1.5 Relevansi Al-Qur'an dalam Penelitian

Dalam konteks penelitian ini, Al-Qur'an dipandang sebagai sumber ayat-ayat tematik yang dapat memberikan dukungan emosional dan spiritual. Sistem rekomendasi berbasis analisis suasana hati memungkinkan pemetaan antara kondisi psikologis pengguna dengan tema ayat yang relevan. Sebagai contoh, pengguna yang sedang sedih dapat diarahkan pada ayat penenang hati (QS. Al-Insyirah [94]:5-6), sedangkan yang marah diarahkan ke ayat pengendalian emosi (QS. Ali Imran [3]:134). Dengan pendekatan ini, penelitian bukan hanya membahas aspek teknis, tetapi juga menghubungkan teknologi dengan bimbingan spiritual berbasis Al-Qur'an [8].

Al-Qur'an memiliki relevansi yang kuat dalam penelitian ini karena berfungsi sebagai sumber utama ayat-ayat tematik yang berkaitan dengan kondisi emosional dan psikologis manusia. Setiap ayat dalam Al-Qur'an mengandung pesan moral, spiritual, dan motivasional yang dapat memberikan ketenangan, penguatan iman, serta arahan perilaku bagi pembacanya. Dengan memanfaatkan ayat-ayat tersebut sebagai basis data, penelitian ini berupaya menghubungkan nilai-nilai spiritual Al-Qur'an dengan pendekatan teknologi modern, sehingga proses pencarian ayat tidak lagi bersifat manual, melainkan dilakukan secara otomatis dan kontekstual berdasarkan suasana hati pengguna [15].

Relevansi Al-Qur'an dalam penelitian ini juga tercermin dalam upaya integrasi antara ilmu agama dan ilmu pengetahuan, khususnya dalam bidang kecerdasan buatan dan pengolahan bahasa alami. Penggunaan teknik *NLP* dan algoritma klasifikasi *SVM* memungkinkan pemetaan antara ekspresi emosi

pengguna dengan tema ayat Al-Qur'an secara sistematis dan terukur. Pendekatan ini menunjukkan bahwa Al-Qur'an tidak hanya relevan sebagai kitab suci dalam aspek ibadah, tetapi juga dapat dijadikan sumber nilai dan inspirasi dalam pengembangan sistem berbasis teknologi yang berorientasi pada kebutuhan spiritual dan emosional manusia di era digital [16].

2.2 Kecerdasan Buatan (*Artificial Intelligence/AI*)

Kecerdasan buatan (*Artificial Intelligence/AI*) mengacu pada kemampuan mesin untuk meniru kecerdasan manusia dalam mempelajari, berpikir, dan mengambil keputusan, ini melibatkan penggunaan algoritma dan teknik komputasi yang kompleks untuk mengolah data, mengenali pola, dan membuat prediksi atau tindakan yang cerdas, contoh penggunaannya adalah *machine learning*, *neural networks*, *natural language processing*, *computer vision*, dan robotika, kecerdasan buatan ini bertujuan untuk meningkatkan efisiensi dan kualitas hidup melalui teknologi cerdas [17]

Kecerdasan buatan (AI) adalah bidang ilmu komputer yang menggunakan simbol daripada angka untuk menyampaikan informasi dan menggunakan metode heuristik atau sejumlah aturan untuk memproses data [27]. Salah satu bidang sains dan teknik yang paling baru adalah kecerdasan buatan, atau AI penelitian tentang kecerdasan buatan dimulai secara resmi setelah Perang Dunia II, dan nama AI pertama kali digunakan pada tahun 1956 [18]

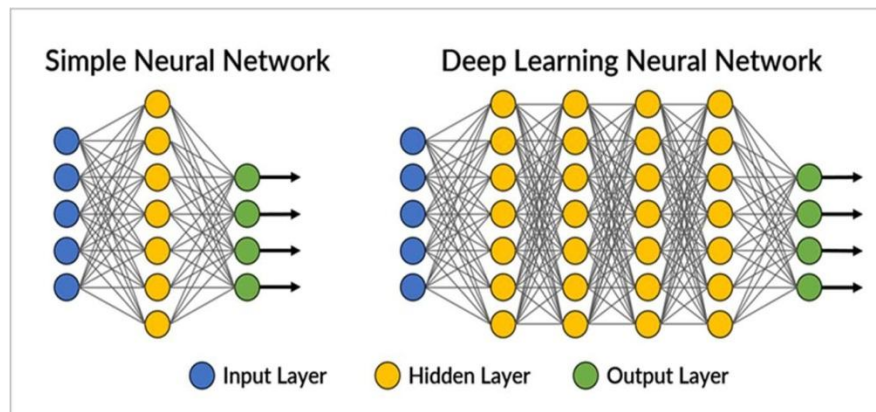
Semakin banyak pengembangan *hardware* dan *software* yang menggunakan teknik kecerdasan buatan menunjukkan betapa berkembangnya kecerdasan buatan saat ini di bidang ilmu komputer [19]. Kemampuan mesin untuk meniru kecerdasan

manusia dalam hal berpikir, belajar, dan membuat keputusan disebut kecerdasan buatan (AI). Ini mencakup penggunaan teknik komputasi yang kompleks dan algoritma untuk mengolah data, mengenali pola, dan membuat prediksi atau tindakan cerdas [20].

2.3 *Deep Learning*

Deep learning merupakan cabang dari algoritma *machine learning* yang memiliki tingkat kompleksitas tinggi. Model dalam *deep learning* dilatih menggunakan sejumlah besar data berlabel dan memanfaatkan arsitektur jaringan saraf (*neural network*) yang mampu mempelajari fitur secara langsung dari data *input*, tanpa memerlukan proses ekstraksi fitur secara manual [21].

Algoritma *deep learning* bekerja dengan memproses data melalui sejumlah lapisan dalam jaringan saraf. Setiap lapisan dalam jaringan tersebut akan menyederhanakan data sebelum mengirimkannya ke lapisan berikutnya. Sementara algoritma *machine learning* tradisional biasanya efektif untuk data terstruktur yang tersusun dalam bentuk baris dan kolom, algoritma tersebut sering mengalami kesulitan ketika dihadapkan pada data tidak terstruktur, seperti gambar, yang membutuhkan proses prapemrosesan yang kompleks. Oleh karena itu, dikembangkanlah algoritma *deep learning* yang dirancang khusus untuk menganalisis dan memproses data visual secara mendalam melalui sejumlah lapisan yang saling terhubung. Lapisan awal berfungsi untuk mengenali detail atau fitur dasar dalam gambar, sedangkan lapisan-lapisan selanjutnya menggabungkan informasi tersebut untuk membentuk representasi yang lebih kompleks dan bermakna [22].



Gambar 2. 1 Struktur *Simple Deep Learning*

Sumber : [22]

Deep Learning dipandang sebagai salah satu terobosan teknologi paling signifikan setelah kemunculan internet, dan kerap digambarkan sebagai inovasi yang berpotensi menggantikan peran manusia dalam berbagai bidang pekerjaan [23].

Deep learning terbagi 3 kategori pendekatan, yaitu:

1. *Supervised learning*

Supervised learning merupakan salah satu cabang dari *machine learning* yang bekerja dengan menggunakan dataset yang telah diberi label. Tujuannya adalah menghasilkan *output* atau solusi yang sesuai berdasarkan data tersebut. Algoritma dalam *Supervised learning* mempelajari pola hubungan antara data dan label yang menyertainya. Beberapa metode yang umum digunakan dalam pendekatan ini antara lain *regresi linier*, *random forest*, dan *support vector machine (SVM)* [21].

2. *UnSupervised learning*

UnSupervised learning merupakan pendekatan deskriptif yang digunakan untuk mengelompokkan atau mengkategorikan data. Tidak seperti *Supervised learning*, metode ini tidak memerlukan dataset yang telah diberi label karena Tujuannya bukan untuk membuat prediksi, melainkan untuk mempelajari struktur dari data yang tersedia. Proses ini dilakukan dengan mengelompokkan data berdasarkan kemiripan atau kedekatan antar data, yang dikenal sebagai *clustering*. Metode ini menganalisis data yang belum berlabel untuk mengidentifikasi pola tersembunyi serta menentukan hubungan atau korelasi di antara data tersebut. Beberapa algoritma yang umum digunakan dalam *UnSupervised learning* antara lain *K-Means*, *Hierarchical Clustering*, *DBSCAN*, dan *Fuzzy C-Means* [21].

3. *Reinforcement learning*

Reinforcement learning merupakan pendekatan pembelajaran mesin yang menggunakan sistem penghargaan (*reward*) dan hukuman (*penalty*) untuk belajar. Algoritma ini berfokus pada bagaimana agen perangkat lunak (*software agent*) mengambil keputusan dalam suatu lingkungan. Tujuan utamanya adalah mempelajari strategi pengambilan keputusan terbaik secara bertahap agar dapat memaksimalkan keberhasilan atau performa dalam situasi nyata [21].

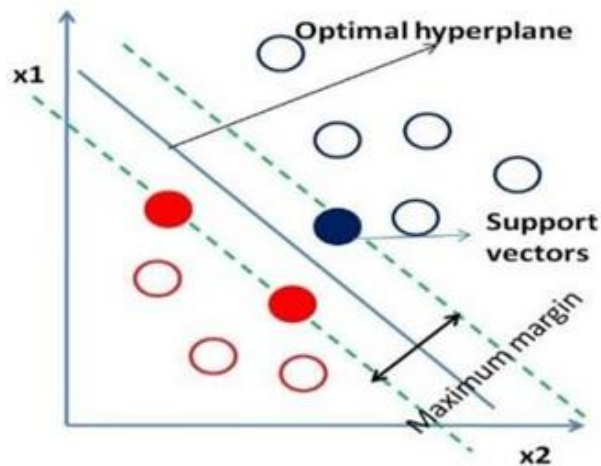
Berdasarkan penjelasan di atas, maka dapat disimpulkan bahwasanya *Deep Learning* merupakan cabang dari *Machine Learning* yang memiliki tingkat kompleksitas tinggi dan menggunakan arsitektur jaringan saraf (*neural network*) untuk mempelajari fitur langsung dari data berlabel tanpa ekstraksi manual.

Teknologi ini dianggap sebagai salah satu inovasi besar setelah internet karena potensinya dalam menggantikan peran manusia di berbagai bidang. *Deep learning* terbagi menjadi tiga pendekatan utama, yaitu *Supervised learning* yang bekerja dengan data berlabel untuk mempelajari hubungan antara *input* dan *output*, *UnSupervised learning* yang menganalisis data tanpa label untuk menemukan pola atau struktur tersembunyi melalui teknik seperti *clustering*, dan *Reinforcement learning* yang menggunakan sistem *reward* dan *penalty* untuk melatih agen dalam mengambil keputusan optimal di lingkungan tertentu.

2.4 *Support Vector Machine (SVM)*

SVM merupakan salah satu algoritma klasifikasi yang dikembangkan oleh Boser, Guyon dan Vapnik yang digunakan untuk memisahkan data berdasarkan 2 kelas berbeda melalui garis optimal hyperplane. *Support Vector Machine (SVM)* adalah metode *learning machine* yang bekerja atas prinsip *Structural Risk Minimization (SRM)* dengan tujuan menemukan *hyperplane* terbaik yang memisahkan dua buah class pada input space. Model algoritma *SVM* merupakan salah satu algoritma dari metode klasifikasi, yang bekerja dengan cara mencari suatu garis (*hyperplane*) untuk memisahkan dua kelompok data. *Hyperplane* dengan pemisah terbaik dapat ditemukan dengan mengukur margin dari *hyperplane* dan mencari titik maksimum. Kernel harus digunakan untuk mencapai keberhasilan banyak algoritma klasifikasi untuk permukaan linier. Dengan demikian dapat diketahui bahwa jenis kernel dapat mempengaruhi hasil klasifikasi yang dilakukan. *Hyperplane* adalah garis pemisah terbaik antara dua kelas. Untuk mencari *hyperplane* dapat dilakukan dengan mencari *margin hyperplane* dan mencari titik

maksimum. *Margin* adalah jarak antara data terdekat di antara dua kelas yang berbeda, yang disebut dengan *support vector* [24].



Gambar 2. 2 Support Vector Machine (SVM)

Sumber : [24]

Pada *SVM* masalah pengklasifikasian dilakukan dengan mencari pemisah *hyperplane* yang optimal antar kelas. *Hyperplane* bergantung dari kasus pelatihan yang meletakkannya pada tepi dari deskriptor kelas yang disebut *support vector*. Prinsip dasar *SVM* adalah *linear classifier*, dan selanjutnya dikembangkan agar dapat bekerja pada masalah *non-linear*, dengan memasukkan konsep *kernel trick* pada ruang kerja berdimensi tinggi [25].

Rumus *SVM* yang sering dipakai adalah sebagai berikut [26]:

1. Polynomial Kernel

Polynomial Kernel merupakan proses yang harus dilakukan untuk mengklasifikasikan dataset training yang sudah normal. Rumus ini dapat di implementasikan menggunakan rumus persamaan 1 sebagai berikut :

$$K(X_i, X_j) = (X_i \cdot X_j + 1)^d \quad \dots\dots\dots(1)$$

2. Radial Bias Function (RBF)

RBF merupakan salah satu rumus fungsi dalam *SVM* yang digunakan untuk klasifikasi dataset yang tidak terpisah secara linier. Pada proses ini *RBF* mempunyai kelebihan tingkat akurasi *training* dan *prediction* yang sangat baik. Rumus ini dapat di implementasikan menggunakan persamaan 2 berikut :

$$K(X_i, X_j) = \exp(-\gamma \|X_i - X_j\|^2) \dots\dots\dots(2)$$

3. Sigmoid Kernel

Sigmoid merupakan proses pengembangan dari jaringan saraf tiruan. Rumus ini dapat diimplementasikan menggunakan persamaan 3 berikut :

$$K(X_i, X_j) = \tanh(\alpha x_1 + X_j + \beta) \dots\dots\dots(3)$$

Keterangan :

K = Kernel

X & Y = *Vector input space*

d = *quadratic*

α = *scalar parameter*

2.5 Dataset

Dataset adalah kumpulan data yang terorganisir dalam format tertentu yang digunakan untuk tujuan analisis, pemodelan, atau pembelajaran mesin (*machine learning*). Dataset dapat berupa data numerik, teks, citra, audio, maupun kombinasi dari berbagai jenis data. Dalam konteks *machine learning*, dataset biasanya terdiri dari dua elemen penting, yaitu fitur (*attributes/variables*) yang menggambarkan karakteristik data, serta label (*class/target*) yang merepresentasikan kategori atau

hasil yang diinginkan. Kejelasan struktur dataset sangat memengaruhi validitas hasil analisis dan performa model yang dibangun [27].

Dalam penelitian ini, dataset memegang peranan penting sebagai fondasi utama dalam proses pelatihan dan pengujian model klasifikasi emosi. Dataset teks emosi berbahasa Indonesia digunakan untuk merepresentasikan berbagai ekspresi suasana hati pengguna, seperti cinta, marah, sedih, senang, dan takut. Keberagaman serta kualitas dataset sangat memengaruhi kemampuan model dalam mengenali pola emosi secara akurat. Oleh karena itu, pemilihan dataset dilakukan secara selektif dengan mempertimbangkan kelengkapan label, kejelasan konteks kalimat, serta kesesuaian bahasa yang digunakan agar hasil klasifikasi dapat mencerminkan kondisi emosional pengguna secara realistis [14].

Selain dataset teks emosi, penelitian ini juga menggunakan dataset ayat Al-Qur'an terjemahan Bahasa Indonesia sebagai sumber rekomendasi. Dataset ini disusun secara terstruktur dan dikelompokkan berdasarkan tema emosi tertentu, sehingga dapat dipetakan dengan hasil klasifikasi yang dihasilkan oleh model SVM. Penggunaan terjemahan resmi memastikan keakuratan makna ayat dan meminimalkan kesalahan interpretasi. Dengan mengombinasikan kedua jenis dataset tersebut, sistem mampu menghubungkan hasil analisis emosi pengguna dengan ayat Al-Qur'an yang relevan secara konsisten dan sistematis, sehingga rekomendasi yang diberikan bersifat kontekstual dan dapat dipertanggungjawabkan secara akademik [15].

2.6 *Machine learning*

Machine learning merupakan cabang dari kecerdasan buatan yang menitikberatkan pada penggunaan algoritma dan metode tertentu untuk melakukan prediksi, mengenali pola, serta melakukan klasifikasi. Beberapa algoritma yang umum digunakan dalam *machine learning* antara lain *neural network*, *decision tree*, *k-nearest neighbor*, *naïve Bayes*, *random forest*, dan lainnya [28].

Machine learning memiliki cabang yang memanfaatkan algoritma jaringan saraf tiruan untuk belajar dan beradaptasi terhadap volume data yang besar. Algoritma jaringan saraf tiruan dalam *deep learning* terinspirasi dari cara kerja otak manusia. Dengan pendekatan ini, mesin dapat mengenali pola dalam data yang tidak terstruktur atau data yang memiliki fitur yang sulit ditentukan secara eksplisit, seperti gambar, teks, audio, dan video. Proses pembelajaran yang dimaksud berbeda dengan cara manusia belajar, karena mesin dianggap "belajar" saat ia dapat memperbarui parameternya, di mana parameter ini secara umum merepresentasikan pengetahuan yang dimiliki oleh mesin tersebut [21].

Machine learning memungkinkan sistem komputer untuk melakukan proses pembelajaran secara otomatis berdasarkan data tanpa harus diprogram secara eksplisit untuk setiap kondisi. Melalui proses pelatihan (*training*), *model machine learning* dapat mengenali pola, hubungan, dan karakteristik tertentu dalam data, kemudian menggunakannya untuk melakukan prediksi atau klasifikasi terhadap data baru. Dalam penelitian ini, machine learning berperan sebagai pendekatan utama untuk mengklasifikasikan suasana hati pengguna berdasarkan teks masukan, di mana algoritma *Support Vector Machine (SVM)* digunakan untuk membangun

model yang mampu memisahkan kategori emosi secara optimal. Penerapan machine learning ini menjadikan sistem lebih adaptif, akurat, dan efisien dalam memberikan rekomendasi ayat Al-Qur'an yang sesuai dengan kondisi emosional pengguna [29].

2.7 *Natural Language Processing (NLP)*

Natural Language Processing (NLP) adalah salah satu cabang dalam kecerdasan buatan (AI) yang berfokus pada interaksi antara komputer dan bahasa manusia. Dalam AI, *NLP* memiliki peran penting karena memungkinkan komputer untuk menganalisis, memahami, dan menghasilkan bahasa alami manusia, baik dalam bentuk teks maupun suara. Secara singkat, *NLP* bertujuan untuk membuat mesin lebih "mengerti" bahasa yang kita gunakan sehari-hari, sehingga komunikasi antara manusia dan mesin menjadi lebih intuitif dan efisien [30].

Natural Language Processing (NLP) adalah sub-bidang kecerdasan buatan yang berfokus pada interaksi antara komputer dan bahasa manusia. Tujuan dari teknologi ini adalah untuk membuat mesin mampu membaca dan menalar dengan bahasa manusia dan secara otomatis memprosesnya [31].

Natural Language Processing (NLP) berperan penting dalam menjembatani komunikasi antara bahasa manusia dan sistem komputer melalui proses analisis dan pemahaman teks secara otomatis. Dalam konteks penelitian ini, *NLP* digunakan untuk mengolah teks ekspresi emosi pengguna agar dapat dipahami oleh algoritma machine learning. Tahapan *NLP* seperti *case folding*, *tokenisasi*, *stopword removal*, dan *stemming* bertujuan untuk mengurangi noise serta mengekstraksi informasi penting dari teks. Dengan penerapan *NLP* yang tepat, representasi teks menjadi

lebih terstruktur dan bermakna, sehingga meningkatkan kinerja algoritma klasifikasi dalam mengenali pola emosi dan mendukung proses rekomendasi ayat Al-Qur'an secara akurat dan kontekstual [14].

2.8 *Google Collab*

Google Colab adalah sebuah *IDE* untuk pemrograman *Python* dimana pemrosesan akan dilakukan oleh *server Google* yang memiliki perangkat keras dengan performa yang tinggi. Dari sisi perangkat lunak, *Google Colab* telah menyediakan hampir sebagian besar pustaka (*library*) yang dibutuhkan. Dalam penelitian ini pustaka yang dibutuhkan adalah Keras, *TensorFlow*, *NumPy*, *Pandas*, dan pendukung lainnya, misalnya untuk pembuatan grafik lewat *Matplotlib*. Bahkan seluruh versi disediakan, misalnya *TensorFlow* versi 1.x maupun 2.x tersedia, begitu juga versi *Python* yang mulai dari versi 2.x hingga 3.x [32].

Google Collab adalah *platform* yang memungkinkan pengguna untuk membuat dan berbagi kode secara *online* dengan mudah dan gratis dengan menggunakan bahasa pemrograman *Python* yang menyediakan fitur built in interpreter [33].

Google Colaboratory, juga dikenal sebagai *Colab*, adalah *platform* komputasi awan gratis yang menyediakan lingkungan *notebook* tanpa *server* interaktif untuk menulis, berbagi, dan mengeksekusi kode. *Colab* beroperasi pada *notebook Jupyter*, yang kompatibel dengan bahasa pemrograman populer seperti *Python*, R, dan Julia. *Colab* menawarkan sumber daya komputasi yang kuat, termasuk akses ke mesin virtual jarak jauh berbasis *Linux* dengan unit pemrosesan grafis (GPU) yang canggih hingga 12 jam. Hal ini menjadikannya *platform* yang

nyaman untuk berbagai aplikasi, termasuk tugas pembelajaran mendalam dan pembelajaran mesin [34].

2.9 *Python*

Python adalah bahasa pemrograman yang fleksibel dan sederhana yang didefinisikan dalam dokumen-dokumennya sebagai. *Python* adalah sebuah bahasa pemrograman dinamis yang sering digunakan dalam pengembangan aplikasi pada berbagai domain. Hal ini memungkinkan suatu program ditulis dalam beberapa pendekatan sekaligus. Misalnya, antar muka grafis dibuat dalam bentuk orientasi objek, sedangkan pemrosesan dalam bentuk fungsional atau prosedural. Bahasa pemrograman *Python* memiliki beberapa fitur yang dapat digunakan oleh pengembang perangkat lunak [35].

Python adalah bahasa pemrograman interpretatif multiguna dengan filosofi perancangan yang berfokus pada tingkat keterbacaan kode. *Python* diklaim sebagai bahasa yang menggabungkan kapabilitas, kemampuan, dengan sintaksis kode yang sangat jelas, dan dilengkapi dengan fungsionalitas pustaka standar yang besar serta komprehensif. *Python* juga didukung oleh komunitas yang besar. *Python* mendukung multi paradigma pemrograman, utamanya; namun tidak dibatasi; pada pemrograman berorientasi objek, pemrograman imperatif, dan pemrograman fungsional [36].

Python digunakan dalam penelitian ini sebagai bahasa pemrograman utama karena memiliki ekosistem pustaka yang lengkap dan mendukung pengembangan machine learning serta pengolahan bahasa alami. Berbagai library seperti NumPy dan Pandas dimanfaatkan untuk pengolahan dan manipulasi data, sedangkan Scikit-

learn digunakan untuk implementasi algoritma Support Vector Machine (SVM). Selain itu, library NLP seperti NLTK dan Sastrawi membantu dalam proses prapemrosesan teks berbahasa Indonesia. Kemudahan sintaks dan dukungan komunitas yang luas menjadikan Python efektif dalam membangun, menguji, dan mengevaluasi sistem klasifikasi emosi serta rekomendasi ayat Al-Qur'an secara efisien dan terstruktur [15].

2.10 Penelitian Terdahulu

Untuk penelitian terdahulu dapat dilihat pada tabel berikut:

Tabel 2. 1 Penelitian terdahulu

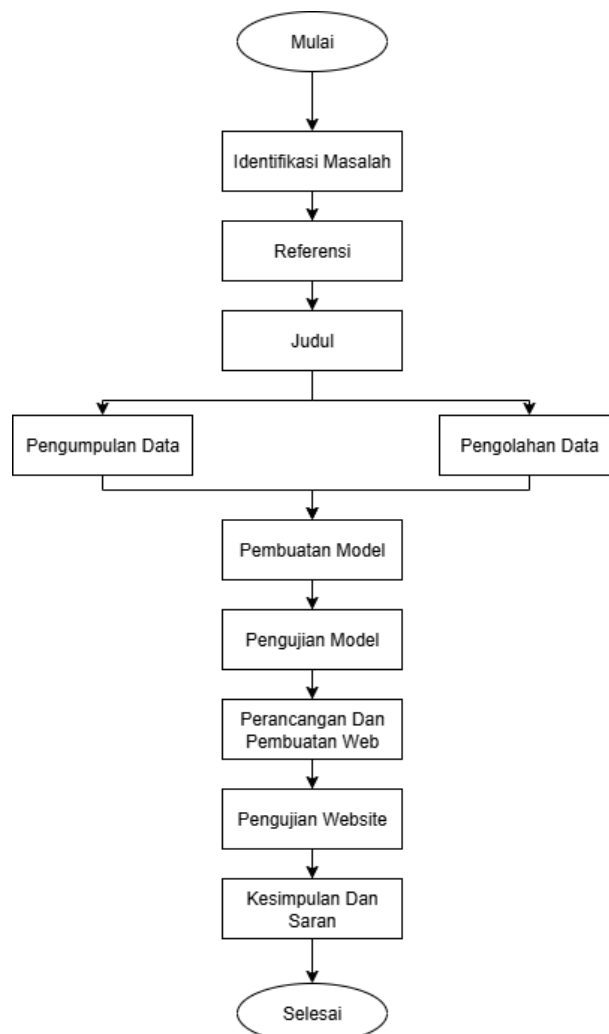
No	Nama Peneliti (Tahun)	Judul Penelitian	Metode	Kesimpulan
1.	Moch. Fauzan, Hartarto Junaedi, Endang Setyati (2022)	Klasifikasi Al-Qur'an Terjemahan Bahasa Indonesia Dengan Menggunakan Algoritma <i>Support Vector Machine (SVM)</i>	Preprocessing teks (<i>tokenizing, stopword removal, stemming</i>), ekstraksi fitur TF-IDF, klasifikasi dengan <i>SVM</i> , uji dengan <i>Chi-Square & PCA</i> untuk <i>feature selection</i> .	<i>SVM</i> tanpa <i>feature selection</i> memberi performa terbaik dengan $AUC \approx 83,3\%$. Artinya <i>SVM</i> cukup efektif mengklasifikasikan ayat ke dalam kategori topik tertentu.
2.	Annisa Azka Pambudi, Kemas Muslim Lhaksmana (2022)	Identifikasi Ayat Sulit dalam Al-Quran Menggunakan <i>SVM</i> , LSTM, dan CNN	Preprocessing teks terjemahan, ekstraksi fitur TF-IDF, klasifikasi menggunakan <i>SVM</i> , LSTM, dan CNN. Menangani data imbalance	<i>SVM</i> menghasilkan akurasi tertinggi ($\approx 68\%$), lebih baik daripada CNN (66%) dan LSTM (65%), sehingga <i>SVM</i> dinilai cocok untuk klasifikasi ayat sulit vs mudah.

			dengan <i>SMOTE</i> & <i>undersampling</i> .	
3.	Novi Amalia Hidayat (2020)	Klasifikasi Teks Hasil Tafsir Ayat Sains dalam Al-Qur'an Menggunakan Algoritma <i>SVM</i>	Korpus tafsir ayat sains, preprocessing teks, representasi fitur (TF-IDF), klasifikasi dengan <i>SVM</i> .	<i>SVM</i> mampu mengidentifikasi ayat tafsir bertema sains secara lebih terstruktur sehingga mempermudah pencarian ayat terkait ilmu pengetahuan.
4.	Fadila Khair (2021)	Analisis Support Vector Machine dalam Pengenalan Bacaan Ayat Al-Qur'an Surah Al-Mu'minun Ayat 1-5	Pengenalan suara menggunakan fitur MFCC dari rekaman bacaan Al-Qur'an, klasifikasi dengan <i>SVM</i> .	<i>SVM</i> dapat mengenali bacaan ayat dengan akurasi cukup baik, meskipun variasi bacaan masih mempengaruhi hasil.
5.	Z. Putriando (2024)	Sistem Rekomendasi Al-Qur'an Berbasis Topik	Kategorisasi ayat berdasarkan topik, perhitungan similarity antar teks untuk rekomendasi ayat.	Sistem dapat merekomendasikan ayat sesuai topik yang dibutuhkan pengguna, menjadi dasar bagi pengembangan sistem rekomendasi yang lebih personal.

BAB III

METODOLOGI PENELITIAN

Tahapan penelitian yang akan dilakukan dan penyelesaian masalah terhadap sistem rekomendasi ayat al-qur'an otomatis berdasarkan analisis suasana hati pengguna menggunakan NLP dan algoritma klasifikasi *SVM*. Adapun tahapan metodologi yang dilakukan selama penelitian dapat dilihat pada Gambar 3.1 yang mana merupakan proses yang dimulai dari studi literatur hingga diperoleh kesimpulan.



Gambar 3. 1 Tahapan Metodologi Penelitian

3.1 Identifikasi Masalah

Identifikasi masalah merupakan tahap awal dalam penelitian ini. Masalah utama yang ditemukan adalah bagaimana cara membangun sebuah sistem rekomendasi ayat Al-Qur'an yang mampu memberikan ayat sesuai dengan kondisi suasana hati (emosi) pengguna secara otomatis.

Selama ini, sebagian besar aplikasi Al-Qur'an digital hanya menyediakan fitur pencarian berdasarkan kata kunci atau tema tertentu, misalnya “sedih, marah, atau cemas”. Hal ini mengharuskan pengguna mengetahui terlebih dahulu kata kunci atau topik yang relevan dengan kondisinya. Kekurangan dari pendekatan tersebut adalah tidak adanya sistem yang mampu memahami teks ekspresi perasaan pengguna, kemudian secara otomatis menampilkan ayat yang sesuai tanpa perlu pencarian manual.

Permasalahan ini menjadi semakin penting karena perkembangan teknologi kecerdasan buatan (*Artificial Intelligence*) memungkinkan penerapan *Natural Language Processing (NLP)* untuk analisis teks, termasuk dalam mengenali suasana hati. Selain itu, *Support Vector Machine (SVM)* dikenal sebagai algoritma yang efektif dalam klasifikasi teks dengan tingkat akurasi yang baik.

3.2 Referensi

Tahap referensi bertujuan untuk memperkuat dasar penelitian dengan melakukan kajian pustaka. Beberapa topik utama yang menjadi dasar teori adalah:

1. Analisis Sentimen Teks

digunakan untuk mengenali suasana hati dari input pengguna, dengan dukungan teori *NLP*. Datasetnya di dapat di *Github link* nya yaitu: <https://github.com/meisaputri21/Indonesian-Twitter-Emotion-Dataset>.

2. *Natural Language Processing (NLP)*

sebagai metode utama dalam pengolahan teks (*case folding*, tokenisasi, *stopword removal*, *stemming*, dan ekstraksi fitur).

3. *Algoritma Support Vector Machine (SVM)*

dipilih karena performanya yang baik pada masalah klasifikasi teks.

4. Sistem Rekomendasi

digunakan untuk memetakan hasil klasifikasi ke ayat-ayat Al-Qur'an sesuai kategori suasana hati.

5. Ayat Al-Qur'an Digital

menggunakan terjemahan resmi dari Kementerian Agama RI sebagai sumber dan di dapat di Kaggle link nya yaitu <https://www.kaggle.com/datasets/ahmadreginald/quran-arabic-indonesian-parallel-corpus>.

Referensi diperoleh dari jurnal internasional, jurnal nasional terindeks SINTA, prosiding seminar, serta buku-buku terkait machine learning, NLP, dan tafsir tematik Al-Qur'an.

3.3 Penentuan Judul

Penentuan judul penelitian dilakukan dengan mempertimbangkan masalah yang ditemukan serta hasil kajian literatur. Judul dipilih agar mencerminkan ruang lingkup penelitian sekaligus inovasi yang ditawarkan.

Judul yang diputuskan adalah: “Sistem Rekomendasi Ayat Al-Qur’an Otomatis Berdasarkan Analisis Suasana Hati Pengguna Menggunakan NLP dan Algoritma Klasifikasi *SVM*.”

Judul ini mengandung beberapa poin penting:

1. Sistem Rekomendasi Ayat Al-Qur’an
menunjukkan fokus penelitian yaitu pengembangan sistem Islami.
2. Analisis Suasana Hati Pengguna
menekankan penggunaan NLP untuk memahami emosi pengguna.
3. NLP dan Algoritma *SVM*
menunjukkan metode yang digunakan dalam penelitian.

3.4 Pengumpulan Data

Data yang digunakan dalam penelitian ini terdiri dari dua jenis:

1. Data Ekspresi Suasana Hati
 - 1) Berupa teks bahasa Indonesia yang mengekspresikan perasaan sedih, marah, atau cemas.
 - 2) Data diperoleh dari dataset *Kaggle*.
2. Data Ayat Al-Qur’an
 - 1) Sumber data utama adalah terjemahan Al-Qur’an resmi dari Kementerian Agama RI.
 - 2) Data diorganisasi berdasarkan tema tertentu, misalnya ayat tentang sedih, marah, atau cemas.
 - 3) Setiap ayat diberi anotasi tema sehingga dapat dipetakan ke kategori suasana hati.

Pengumpulan data dilakukan secara sistematis dan disimpan dalam format digital (*CSV* dan *JSON*) agar mudah diproses.

3.5 Pengolahan Data

Pengolahan data merupakan tahap penting sebelum data digunakan untuk pelatihan model. Tahap-tahap pengolahan data meliputi:

1. *Case Folding*

mengubah semua huruf menjadi huruf kecil agar konsisten.

2. Tokenisasi

memecah teks kalimat menjadi unit kata.

3. *Stopword Removal*

menghapus kata-kata yang tidak memiliki makna penting, misalnya “yang”, “dan”, “di”.

4. *Stemming/Lemmatization*

mengembalikan kata ke bentuk dasar, misalnya “berlari”, “lari-lari” → “lari”.

5. Normalisasi Teks

mengubah kata tidak baku ke bentuk baku, misalnya “gk” → “tidak”.

6. Ekstraksi Fitur

hasil preprocessing dikonversi ke bentuk numerik menggunakan metode *TF-IDF* (*Term Frequency – Inverse Document Frequency*), sehingga bisa diproses oleh algoritma *SVM*.

3.6 Pembuatan Model

Pembuatan model dilakukan dengan algoritma *Support Vector Machine* (SVM) karena dikenal efektif untuk klasifikasi teks dengan margin maksimal.

Tahapan pembuatan model meliputi:

1. Split Dataset

Pada tahap pembagian data (split dataset), seluruh dataset yang berjumlah 1.000 data teks emosi dibagi menjadi dua bagian, yaitu 800 data (80%) sebagai data latih (training) dan 200 data (20%) sebagai data uji (testing). Data latih digunakan untuk membangun dan melatih model Support Vector Machine (SVM) agar mampu mengenali pola emosi dalam teks, sedangkan data uji digunakan untuk mengevaluasi performa model terhadap data yang belum pernah dilihat sebelumnya. Pembagian ini dilakukan secara acak untuk memastikan distribusi kelas emosi tetap seimbang pada kedua kelompok data sehingga hasil evaluasi lebih objektif dan representatif.

2. Training

model dilatih dengan data latih menggunakan kernel *Radial Basis Function* (RBF).

3. Optimization

tuning hyperparameter (C dan gamma) untuk mendapatkan hasil terbaik.

4. Model Saving

model yang sudah terlatih disimpan agar dapat digunakan kembali dalam sistem web.

3.7 Pengujian Model

Pengujian model dilakukan untuk mengevaluasi performa *SVM* dalam mengklasifikasikan suasana hati. Evaluasi menggunakan confusion matrix dan metrik berikut:

1. Akurasi

proporsi prediksi benar terhadap seluruh data uji.

2. Presisi

ketepatan model dalam memprediksi label tertentu.

3. *Recall*

kemampuan model menemukan semua data relevan.

4. *F1-Score*

keseimbangan antara presisi dan *recall*.

Pengujian dilakukan menggunakan data uji (*testing set*) yang terpisah dari data latih agar hasil objektif.

3.8 Perancangan dan Pembuatan Web

Setelah model selesai dibangun, sistem berbasis web dirancang untuk memudahkan akses pengguna. Tahapan meliputi:

1. Perancangan Antarmuka (UI/UX)

desain dibuat sederhana, responsif, dan ramah pengguna dengan menu input teks, hasil klasifikasi, serta rekomendasi ayat.

2. Integrasi *Backend*

menghubungkan input pengguna dengan model *SVM*.

3. Pemetaan Ayat

hasil klasifikasi suasana hati dipetakan ke tema ayat tertentu berdasarkan aturan (*rule-based*).

4. Implementasi Teknologi

menggunakan bahasa pemrograman *Python (Flask/Django/Streamlit)* serta database *CSV/MySQL* untuk penyimpanan data.

3.9 Pengujian Website

Pengujian sistem *web* dilakukan untuk memastikan sistem berfungsi sesuai kebutuhan. Pengujian meliputi:

1. *Functional Testing*

memastikan setiap fitur berjalan, seperti input teks, klasifikasi suasana hati, dan output ayat Al-Qur'an.

2. *Performance Testing*

mengukur kecepatan respon sistem dalam memproses input dan menampilkan hasil.

3. *User Acceptance Test (UAT)*

melibatkan pengguna (misalnya mahasiswa atau masyarakat umum) untuk memberikan *feedback* apakah ayat yang direkomendasikan sesuai dengan suasana hati yang diinputkan.

3.10 Kesimpulan dan Saran

Tahap terakhir adalah menyusun kesimpulan dari hasil penelitian. Kesimpulan berisi sejauh mana sistem dapat:

1. Mengenali suasana hati pengguna berdasarkan teks input.
2. Memberikan rekomendasi ayat yang sesuai dengan kondisi emosional.

3. Membuktikan bahwa *SVM* dengan preprocessing *NLP* mampu menghasilkan akurasi tinggi.

Selain itu, diberikan saran untuk penelitian berikutnya, misalnya:

1. Menambahkan kategori emosi yang lebih detail (misalnya takut, marah, sedih, bahagia).
2. Menggunakan algoritma deep learning seperti *LSTM* atau *BERT* untuk hasil yang lebih baik.
3. Mengembangkan sistem ke *platform mobile* agar lebih mudah digunakan.