

BAB II

LANDASAN TEORI

2.1 Tinjauan Pustaka

Kecerdasan buatan, juga dikenal sebagai *artificial intelligence* (AI) adalah sistem komputer yang diprogram untuk melakukan hal-hal yang serupa dengan manusia dan telah berkembang dengan sangat cepat sehingga digunakan dalam banyak bidang. Cabang ilmu kecerdasan buatan yang disebut *natural language processing* (NLP), juga dikenal sebagai pengolahan bahasa alami, berfokus pada bagaimana manusia dan komputer berinteraksi dengan perantara pengolahan bahasa natural. Bahasa natural adalah bahasa yang biasa digunakan orang untuk berbicara satu sama lain. Bahasa manusia harus diproses dan dipahami sebelum komputer dapat memahami maksud orang. *Chatbot*, aplikasi yang membuat pengguna merasa seperti mereka berbicara dengan komputer, adalah salah satu contoh aplikasi pengolahan bahasa alami. Kecerdasan buatan dan khususnya pembelajaran mesin dapat digunakan dalam pengembangan solusi prediksi, deteksi, dan perawatan untuk penanganan kesehatan mental (D'Alfonso, 2020).

Kecerdasan buatan yang menerapkan pengolahan bahasa alami pada *chatbot*, telah diterapkan dalam berbagai bidang kehidupan manusia saat ini. Penelitian di bidang ekonomi menyelidiki merek ritel busana mewah, yang menggunakan *chatbot* untuk menyediakan layanan elektronik. Data pelanggan digunakan dalam penelitian untuk mengukur persepsi pelanggan tentang pemecahan masalah, penyesuaian, ketertarikan, interaksi, dan tren. Hasilnya menunjukkan bahwa *chatbot* dapat membantu membangun hubungan yang lebih baik dengan pelanggan. *Chatbot* adalah strategi pemasaran yang dapat meningkatkan kepuasan pelanggan karena pelanggan kaya mengharapkan pengalaman unik (Chung dkk., 2020).

Pada bidang kesehatan sendiri penerapan *chatbot* sudah sangat banyak diterapkan. Dalam penelitian ini, *chatbot* medis dapat mendiagnosis penyakit dan memberikan detail tentang kondisi pasien sebelum pasien berkonsultasi dengan psikiater atau psikolog. Tujuannya adalah untuk mengurangi biaya perawatan, membuat pengetahuan medis lebih mudah diakses, dan menyediakan referensi

medis. Ketika *chatbot* dapat mendiagnosis penyakit dan memberikan informasi, mereka sangat membantu pengguna. Bot akan memeriksa masukan dan membuat diagnosis yang disesuaikan berdasarkan gejala. Akibatnya, orang dapat mengetahui kondisi kesehatan mereka dan mendapatkan perlindungan yang tepat. (Divya dkk., 2018).

Studi menunjukkan bahwa forum interaksi sangat penting untuk menjawab pertanyaan tentang kesehatan pribadi sebelum berkonsultasi dengan dokter. *Chatbot* dibuat untuk membantu pencari kesehatan bekerja lebih baik dan membutuhkan orang yang dapat diandalkan untuk berkomunikasi. Bagian dari proses desain sistem adalah melakukan analisis permintaan pengguna dan memastikan bahwa permintaan tersebut sesuai dengan *dataset* saat ini. Jika keluhan ada dalam *dataset*, jawaban akan diberikan sesuai dengannya; jika tidak, kontak spesialis akan diberikan (Kumar dan Keerthana, 2016).

Penerapan *chatbot* pada bidang kesehatan mental secara umum, telah dikembangkan oleh berbagai peneliti. Dengan menggunakan metode kajian empiris terhadap penelitian-penelitian terkait, *review literature* membahas bagaimana *chatbot* dalam bidang kesehatan mental digunakan oleh pengguna. Hasilnya menunjukkan bahwa penelitian kesehatan mental tentang *chatbot* sebagian besar berbasis aturan dan ditujukan untuk mengobati autisme dan depresi. Bidang ini membutuhkan penelitian lebih lanjut. Selain itu, penyedia layanan kesehatan harus membantu calon pasien mendapatkan dukungan kesehatan mental (Abd-Alrazaq dkk., 2019).

Selanjutnya, terkait penggunaan *chatbot* untuk memberikan dampak positif secara psikologis pada orang usia muda setelah pengobatan kanker dengan tujuan untuk mengurangi tekanan psikososial. Hasil yang didapatkan menunjukkan bahwa *chatbot* yang dirancang dapat diterima untuk memberikan dampak positif secara psikologis pada orang yang dirawat karena kanker (Greer dkk., 2019).

Sebuah *chatbot* diciptakan oleh Park dkk. pada tahun 2019 untuk memberikan motivasi dan membantu siswa mengatasi stres. Sebanyak tiga puluh siswa diwawancarai melalui aplikasi Bonobot, dan metode Braun dan Clarke digunakan untuk menganalisis laporan stres mereka tentang kehidupan sekolah.

Penelitiannya menemukan bahwa pernyataan yang tepat pada *chatbot* dapat membantu orang berpikir tentang diri mereka sendiri, mempelajari masalah mereka, dan mengubah perilaku mereka dengan cara yang lebih baik, seperti yang dilakukan dalam wawancara motivasi (Park dkk., 2019).

Ho, dkk, pada tahun 2018 melakukan perbandingan antara manusia dengan *chatbot*. Studi ini membandingkan rasa emosional, relasional, dan psikologis antara percakapan manusia dan *chatbot*. Studi ini juga mempelajari bagaimana peran manusia dalam berinteraksi dengan teknologi seperti *chatbot* untuk mengungkapkan informasi pribadi. Hasilnya menunjukkan bahwa percakapan antara manusia dan *chatbot* sama-sama menghasilkan manfaat emosional, relasional, dan psikologis. Hasil ini juga mendukung hipotesa Ho bahwa orang akan berinteraksi secara psikologis dengan robot (Ho dkk., 2018).

Pada tahun 2020, penelitian telah dilakukan oleh Tran dkk. dengan tujuan untuk memahami, mendeteksi, dan mengidentifikasi hubungan antara ucapan manusia dan sistem *chatbot*. Tujuan penelitian ini adalah untuk menyarankan kerangka kerja yang dapat dimodelkan untuk digunakan dalam sistem *chatbot* untuk mempelajari fitur baru yang berguna. Di Vietnam, *chatbot e-commerce* menggunakan kerangka tersebut untuk membantu merek ritel berinteraksi dengan pelanggan dengan lebih baik. Hasilnya menunjukkan bahwa kerangka tersebut sangat menjanjikan untuk membantu bot memahami lebih baik apa yang dikatakan pengguna (Tran dan Luong, 2020).

Rukun pada tahun 2017, menyimpulkan bahwa penerapan fitur *chatbot* pada sebuah aplikasi yang memerlukan tambahan sistem berbasis sistem pakar akan membantu memberikan solusi dari permasalahan yang kompleks. Dalam penelitian ini, sebuah aplikasi pakar menggunakan metode *forward chaining* untuk mendeteksi gangguan saluran cerna pada balita. Metode *forward chaining* dapat membantu menyelesaikan suatu permasalahan dari indeks gejala awal hingga bisa menemukan informasi berarti sesuai dengan aturan dari suatu *rules* yang telah ditentukan (Rukun dkk., 2017).

Berdasarkan penelitian yang dilakukan dalam membangun sistem informasi konsultasi kesehatan mental menggunakan pengolahan bahasa alami berbasis

sistem pakar dengan metode *forward chaining*, banyak penelitian sebelumnya yang telah membahas dan mengkaji terkait penerapan *chatbot* pada bidang kesehatan mental dan memiliki dampak yang positif yang dirasakan untuk pengguna. Penggunaan *chatbot* sendiri memiliki pengaruh yang cukup baik pada psikologis pengguna. Selain itu, menerapkan sistem pakar dengan metode *forward chaining* dapat membantu diagnosis kesehatan mental pengguna sistem informasi konsultasi dengan banyak data yang digunakan.

2.2 Dasar Teori

2.2.1 E-Mental Health

Kesehatan mental (*mental health*) adalah keadaan sejahtera di mana seseorang dapat memanfaatkan kapasitasnya, memulihkan diri dari tekanan jadwal hari demi hari, bermanfaat dan berkontribusi kepada masyarakat (World Health Organization, 2004). Beberapa contoh gangguan kesehatan mental di antaranya adalah depresi, kecemasan, dan stres. *E-health* sendiri adalah teknologi internet yang terkait yang bertujuan untuk meningkatkan perawatan kesehatan dengan membangun lebih banyak jalur komunikasi antara pasien dan penyedia layanan. Penerapan *e-health* ada di berbagai bidang kesehatan, sedangkan pengertian *e-mental health* secara khusus sendiri mengacu pada penggunaan internet dan teknologi terkait untuk pemberian atau peningkatan informasi, pelayanan dan perawatan kesehatan mental (Blaauw dan Emerencia, 2016; Florescu dan Florescu, 2020). Penggunaan *e-mental health* menunjukkan bukti kuat yang efektif untuk digunakan dalam pengobatan depresi, kecemasan, dan stres di mana banyak sekali peneliti di dunia yang telah mengkaji dan menganalisis berbagai persepsi dari dampak dan manfaat yang dialami pengguna *e-mental health*, sehingga dapat disimpulkan bahwasanya penggunaan *e-mental health* sangat efektif untuk mempermudah pasien gangguan mental secara efisien dan mudah (Andersson dan Hedman, 2013; Andrews dan Williams, 2015; Ren dkk., 2020).

Berikut ini adalah manfaat dari adanya implementasi *e-mental health* yang dirasakan pengguna (Royal Australian College of General Practitioners, 2015):

1. Nyaman dan fleksibel karena dapat diakses kapan saja dan di mana saja, dan materi dapat ditinjau sesering mungkin.
2. Biaya layanan yang rendah atau tidak ada sama sekali untuk pasien, banyak intervensi gratis, meskipun biaya terkait seperti data seluler atau batas pengunduhan internet harus dipertimbangkan.
3. Mengisi celah layanan, dapat memberikan pengenalan terapi, atau alternatif bagi orang yang tidak menyukai pengobatan tatap muka.
4. Menghemat waktu praktisi, memungkinkan profesional kesehatan mental untuk fokus pada pasien yang pendekatan kesehatan mentalnya tidak sesuai, sehingga mengurangi daftar tunggu.
5. Hemat biaya untuk sistem kesehatan, tidak mahal untuk diberikan, dan dapat disebarluaskan ke populasi besar.
6. Mudah diakses, dapat memberikan dukungan kepada pasien yang tidak dapat menghadiri perawatan tatap muka.

Proses mengidentifikasi bukti kandidat ideal untuk *e-mental health*, pasien yang paling cocok untuk *e-mental health* adalah untuk orang yang memiliki peningkatan risiko gangguan kesehatan mental, atau yang sudah mengalami gejala gangguan kesehatan mental ringan hingga sedang. Ada sedikit bukti tentang penggunaan intervensi kesehatan mental elektronik dengan orang yang memiliki gangguan kesehatan mental yang kompleks atau parah, gangguan kepribadian komorbid, ketergantungan zat, atau orang yang memiliki risiko tinggi untuk melukai diri sendiri atau bunuh diri dan membutuhkan manajemen klinis segera. Pada penelitian ini difokuskan pada kendala kesehatan mental depresi, kecemasan, dan stres berdasarkan data gangguan mental yang umumnya dialami di seluruh dunia. Sasaran pengguna sistem informasi konsultasi kesehatan mental ditujukan pada masyarakat umum dengan usia ≥ 15 tahun dengan indikasi atau mengalami depresi, kecemasan, dan stres. Sesuai dengan rentang usia yang umum untuk gangguan kesehatan mental (Pusat Data dan Informasi Kementerian Kesehatan RI, 2019).

2.2.2 Kecerdasan Buatan

Kecerdasan buatan atau *artificial intelligence* adalah sistem yang dirancang untuk melakukan tugas-tugas yang dilakukan oleh manusia, seperti menyelesaikan masalah, membuat keputusan, berpikir, dan belajar (Kusumadewi, 2003; McCarthy dkk., 2006; Rich dan Knight, 1991). Ketika kecerdasan buatan dikembangkan secara positif, itu dapat membantu semua aktivitas manusia. Kecerdasan buatan mencakup sistem pakar, pengolahan bahasa alami, pengenalan ucapan, robotika dan sistem sensor, penglihatan komputer, komputer pintar dan instruksi, dan bermain *games* (Kamble dan Shah, 2018).

Ada empat kategori kecerdasan buatan. Pertama, berpikir secara manusiawi, bahwa teknologi kecerdasan buatan mencerminkan cara manusia berpikir; kedua, bertindak secara manusiawi, bahwa kecerdasan buatan mampu bertindak secara manusiawi; ketiga, berpikir secara logis, bahwa kecerdasan buatan mampu menyelesaikan masalah secara logis; dan keempat, berpikir secara rasional, bahwa kecerdasan buatan dibuat untuk mencapai tujuan melalui penggunaan sistem. Diharapkan kecerdasan buatan akan membantu proses penyelesaian masalah manusia (Russell dan Norvig, 2020).

Berikut ini merupakan definisi dari kelebihan kecerdasan buatan, antara lain (Arsenijevic dan Jovic, 2019):

1. Kecerdasan buatan lebih permanen karena tidak berubah selama sistem dan program tidak berubah.
2. Lebih mudah untuk menyebarkan karena informasi yang tersimpan dalam sistem komputer masih dapat digunakan bahkan setelah komputer pindah.
3. Lebih murah karena kecerdasan buatan lebih konsisten dalam melakukan hal-hal dalam jangka waktu yang lama.
4. Dapat didokumentasikan karena kecerdasan buatan dapat melacak setiap tindakan melalui penggunaan komputer.

2.2.3 Pengolahan Bahasa Alami

Pengolahan bahasa alami atau *natural language processing* adalah cabang dari ilmu kecerdasan buatan. Proses pengolahan bahasa alami dapat memudahkan komunikasi manusia-mesin atau memahami dan memberikan luaran bahasa

manusia (Liddy, 2001; Shruthi dan Swamy, 2020). Proses pengolahan bahasa alami memodelkan pengetahuan bahasa dari berbagai aspek, termasuk hubungan antara kata-kata dalam suatu kalimat, serta konteks kata-kata dalam kalimat (Cai, 2021).

Tujuan proses pengolahan bahasa alami adalah untuk memberikan tanggapan yang tepat berdasarkan pemahaman mesin tentang makna bahasa manusia (Pustejovsky dan Stubbs, 2013; Shruthi dan Swamy, 2020). Proses pengolahan bahasa alami bertujuan untuk menyelesaikan masalah memahami bahasa alami manusia, termasuk aturan gramatika dan semantiknya, dan mengubahnya menjadi representasi formal yang dapat diproses oleh komputer (Raharjana dkk., 2021).

Ada beberapa masalah dalam menggunakan pengolahan bahasa alami untuk memahami bahasa manusia (Trofin dkk., 2019):

1. Penandaan kelas kata (*part-of-speech tagging*). Sulit untuk menentukan kelas kata apa yang termasuk dalam teks (seperti kata benda, kata kerja, kata sifat, dll.) karena pengelasan kata sangat bergantung pada konteks di mana kata tersebut digunakan.
2. Segmentasi teks (*text segmentation*). Sulit untuk mengidentifikasi segmentasi karena bahasa yang tidak memiliki pembatas kata khusus (seperti Mandarin, Jepang, dan Thailand) dan bahasa lisan yang kadang-kadang membaurkan bunyi antar kata.
3. Disambiguasi makna kata (*word sense disambiguation*). Banyak kata memiliki lebih dari satu arti, baik dalam bentuk homonim (memiliki arti yang berbeda dan tidak terkait, seperti "bisa" yang berarti "dapat" dan "racun") maupun polisemi (memiliki arti yang berbeda tetapi terkait, seperti "ragu" yang berarti "bimbang" dan "sangsi"). Hanya orang yang dapat memahami apa yang digunakan.
4. Ambiguitas sintaksis (*syntactic ambiguity*). Dalam setiap bahasa, ada berbagai jenis struktur kalimat yang dapat digunakan. Biasanya, untuk memilih struktur yang paling cocok, Anda harus menggabungkan informasi kontekstual dan semantik.

5. Masukan yang tak sempurna atau tak teratur (*imperfect or irregular input*). Pemrosesan bahasa alami lebih sulit karena aksen dan kesalahan ejaan dan gramatikal dalam bahasa lisan dan tulisan.
6. Tindak tutur (*speech act*). Kadang-kadang, maksud penutur atau penulis tidak dapat dijelaskan dengan tepat dengan struktur kalimat saja. Dalam beberapa kasus, maksud yang diinginkan ditentukan oleh konteks dan gaya bahasa yang digunakan.

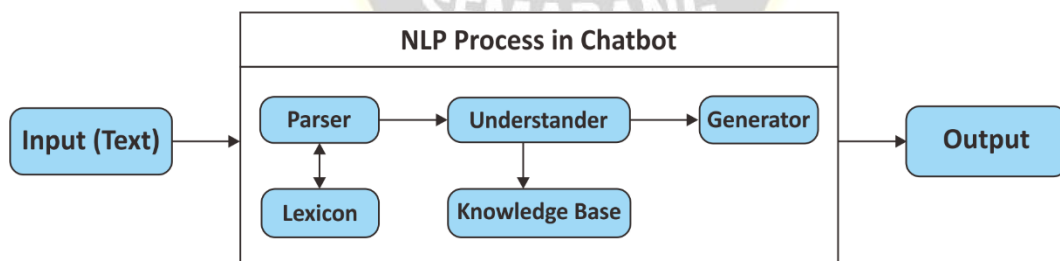
Selain masalah saat ini, proses pengolahan bahasa alami telah digunakan untuk berbagai jenis pekerjaan yang telah dilakukan oleh manusia sebelumnya. Beberapa pekerjaan yang paling populer adalah:

1. Pemerolehan informasi (*information retrieval*). Pencarian dokumen yang relevan, pencarian informasi khusus di dalam dokumen, dan pembuatan metadata adalah semua bagian dari ini.
2. Penjawaban pertanyaan (*question answering*). Secara otomatis memberikan dan menjawab pertanyaan dalam bahasa alami.
3. Perangkuman otomatis (*automatic summarization*). Hal ini melibatkan menggunakan program komputer untuk membuat versi singkat dari dokumen.
4. Penerjemahan mesin (*machine translation*). Proses penerjemahan otomatis dari satu bahasa alami ke bahasa yang berbeda.
5. Pengenalan ucapan (*speech recognition*). Pengenalan bahasa lisan berarti masukan yang dikenali oleh mesin; contohnya, komputer dapat menggunakan bahasa lisan untuk menulis atau memberi perintah dengan bahasa lisan manusia.
6. Sintesis bunyi (*speech synthesis*). pengenalan bahasa tulis menjadi bahasa lisan, yang merupakan kebalikan dari pengenalan bahasa dalam bahasa lisan.
7. Pengenalan karakter optis (*optical character recognition*). Menggunakan pemindai untuk mengubah tulisan tangan atau teks tercetak menjadi dokumen yang dapat dikenali mesin.

8. Analisis sentimen (*sentiment analysis*). Proses mengekstrak informasi dari sumber data teks untuk menentukan sudut pandang yang positif atau negatif tentang suatu objek. Biasanya digunakan untuk menentukan persepsi publik tentang bisnis atau barang.

Secara umum, ada dua jenis aplikasi dalam bidang proses pengolahan bahasa yang berbeda: aplikasi berbasis teks dan aplikasi berbasis dialog. Aplikasi berbasis teks biasanya menggunakan bahasa lisan atau pengenalan suara untuk melakukan proses terhadap teks tertulis, seperti dokumen, *e-mail*, buku, dan sebagainya. Aplikasi berbasis dialog biasanya menggunakan bahasa lisan atau pengenalan suara untuk melakukan interaksi dialog dengan mengetikkan teks pertanyaan melalui *keyboard*. Pada penelitian ini fokus penerapan NLP pada sistem informasi konsultasi kesehatan mental adalah pada fungsi penjawaban pertanyaan dengan fitur *chatbot* yang berbasis *text*. NLP sendiri merupakan salah satu algoritma yang paling populer digunakan untuk pengembangan *chatbot* (Vijayaraghavan dkk., 2020).

Pengolahan bahasa alami telah digunakan untuk memenuhi semua kebutuhan manusia di berbagai bidang kehidupan karena lebih mudah untuk menjadi antar muka komputer daripada mempelajari bahasa perintah-perintah komputer. Tahapan proses dari lima elemen utama dalam pengolahan bahasa alami ditunjukkan pada Gambar 2.1.



Gambar 2.1 Elemen *Chatbot* (Nithuna dan Laseena, 2020)

Parser, *lexicon*, *understander*, *knowledge base*, dan *generator* adalah komponen pengolahan bahasa alami, seperti yang digambarkan dalam Gambar 2.1. *Parser* adalah komponen yang mengidentifikasi setiap kata, dan *lexicon* adalah kumpulan kata yang dikenali program. *Understander* adalah bagian yang menentukan arti sebuah kalimat. *Knowledge base* adalah sumber pengetahuan yang

terdiri dari kata dan frasa. *Generator* adalah *output* yang dibuat dari masukan yang telah diolah. (Nithuna dan Laseena, 2020).

2.2.4 Sistem Pakar

Expert system atau sistem pakar merupakan sub bidang dari kajian ilmu pengetahuan teknologi yaitu kecerdasan buatan, dengan dicirikan oleh adanya pemisah yang jelas antara representasi pengetahuan tentang area masalah khusus *knowledge base* (basis pengetahuan) dan pemrosesan pengetahuan bidang lainnya. Akibatnya, pengetahuan tidak dikodekan dalam algoritma konvensional, melainkan disimpan secara terpisah dari algoritma pemrosesan dalam basis pengetahuan. Kebanyakan dari sistem pakar memiliki standar yang terdiri dari komponen dialog antarmuka (*user interface*), pengetahuan khusus kasus (*human expert*), pengetahuan berbasis aturan (*knowledge base*), komponen pemrosesan pengetahuan (*inference engine*), komponen penjelasan (*explanation facility*) dan komponen akuisisi pengetahuan (*knowledge acquisition*) (Müller dkk., 2021). Sistem pakar yang baik berfungsi untuk menyelesaikan masalah dengan meniru pekerjaan ahli (Kusumadewi, 2003).

Sistem pakar memiliki arsitektur, komponen, dan fungsi yang dapat meningkatkan kemampuan pengambilan keputusan secara menyeluruh untuk mendiagnosis penyakit. Sistem pakar menggunakan kemampuan manusia untuk menggambarkan masalah yang membutuhkan kecerdasan manusia. Sistem pakar juga memberikan pengetahuan ahli dalam bentuk aturan keputusan. Pada saat ini banyak pengimplementasian algoritma sistem pakar dalam bidang medis, yang mana telah dikembangkan untuk membantu diagnosis atau pengobatan penyakit (Hamedan dkk., 2020).

2.2.5 Forward Chaining

Chaining adalah istilah yang digunakan untuk menggambarkan sebuah rantai yang berfungsi untuk menghubungkan masalah dengan solusi. Sebuah rantai dapat dilihat untuk melewati atau menyeberangi suatu masalah untuk mendapatkan solusi. Tahapan di mana dapat menggambarkan pengertian dari *forward chaining* adalah dengan memikirkan fakta dapat berupa data, bukti, temuan atau pengetahuan

yang mengarah pada sebuah kesimpulan berupa kepentingan, penjelasan, atau diagnosis bahwasanya fakta itu ada. *Forward chaining* juga dikenal sebagai penalaran dari bawah ke atas karena penalaran dari tingkat yang lebih rendah ke tingkat atas menghasilkan kesimpulan yang didasarkan pada fakta yang kuat dan kemudian dimasukkan ke dalam basis data. Penalaran dari bawah ke atas adalah bagian dari sistem pakar. Fakta yang kuat sendiri adalah dasar dari pengetahuan berbasis paradigma karena tidak dapat dipecah/diurai menjadi yang sebuah makna terkecil dengan satuan yang memiliki arti (Shofi dkk., 2016).

Kegunaan dari metode *forward chaining* terjadi jika sebuah pohon keputusan menjadi melebar dan tidak mendalam (*deep position*). Pada dasarnya *forward chaining* metode ini membuat sebuah pencarian yang meluas dan untuk lebih mudah, serta metode ini sangat bagus jika mencari sebuah kesimpulan tahap demi tahap (*step by step*). *Forward chaining* dapat dianggap sebagai strategi yang berasal dari sejumlah fakta yang ingin diketahui secara detail, sehingga meminimalkan adanya praduga yang tidak sesuai dengan fakta yang diberikan. Dengan demikian, dapat disimpulkan bahwa metode ini adalah salah satu pendekatan yang baik untuk mengambil keputusan yang diharuskan keakuratan faktanya (Shofi dkk., 2016). Metode penalaran *forward chaining* digunakan untuk menentukan *output* dari pertanyaan yang memiliki beberapa kondisi yang dituangkan dalam aturan, berdasarkan *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) dan Pedoman Penggolongan Diagnosis Gangguan Jiwa (PPDGJ III). Pengguna akan diberikan pertanyaan, dan mereka harus menjawabnya sesuai dengan aturan yang ada. Setelah memproses *input* dan *chatbot* akan menggunakan *inference engine forward chaining* untuk menghasilkan *output*.

2.2.6 Chatbot

Chatbot adalah program dialog yang mengumpulkan informasi dari kata kunci atau frasa (Jurafsky dan Martin, 2020). *Chatbot* sendiri adalah aplikasi yang berkomunikasi dengan mesin dan membantu orang mencari dan mengambil informasi, menjawab pertanyaan, membuat keputusan berdasarkan fakta, membantu belanja, menjadi mitra bahasa, atau bekerja di bidang pendidikan (Nithuna dan Laseena, 2020). Sejarah *chatbot* sangat terkait erat dengan psikologi,

ELIZA merupakan *chatbot* psikolog pertama yang dikembangkan oleh Joseph Weizenbaum, seorang profesor ahli MIT (Massachusetts Institute of Technology) pada tahun 1966 (Weizenbaum, 1966). Pada awal waktu dikembangkan, *chatbot* psikologi kesehatan mental masih sangat sederhana, hingga kini sudah terdapat tiga *chatbot* kesehatan mental, yaitu Woebot, Wysa dan Tess yang unggul dan berkembang sangat cepat. Sesuai dengan perkembangan zaman dan ilmu pengetahuan teruslah dibuat inovasi dengan dibantunya kecerdasan buatan, namun daya tarik dari sistem *chatbot* sendiri masih mempertahankan kedudukannya dalam dunia kecerdasan buatan. *Chatbot* semakin populer karena lebih mudah dan menguntungkan dengan perangkat lunak. (Klopfenstein dkk., 2017).

Chatbot terdiri dari tiga komponen: *responder*, *classifier*, dan *graph master*. Bagian yang menghubungkan bot dengan pengguna adalah *responder*. Mentransfer data dari pengguna ke *classifier* dan memantau *input* dan *output* adalah salah satu fungsi *responder*. *Classifier* menyaring dan menormalkan *input* dan melakukan segmentasi kalimat yang telah dimasukkan. *Graph master* mencocokkan pola *input* dengan fakta yang telah dibuat sebelumnya (Abdul-Kader dan Woods, 2015). *Chatbot* mengumpulkan *input* pengguna dan menghasilkan *output* seperti suara, teks, gambar, dan video, memberi pengguna kesan seperti mereka sedang berbicara dengan manusia (Chen dkk., 2021). *Chatbot* akan mencocokkan kalimat yang dimasukkan pengguna dengan pola pengetahuan yang sudah ada. Pengetahuan ini dikumpulkan dari berbagai sumber dan digunakan untuk membentuk pola komunikasi antara pengguna dan *chatbot*. *Chatbot* diagnosis kesehatan mental akan diimplementasikan pada sistem informasi yang penulis kembangkan sesuai dengan kebutuhan pengguna.

2.2.7 Text Mining

Text mining juga dikenal sebagai penambangan teks, adalah suatu metode untuk mengidentifikasi pola dan mengumpulkan informasi dari data teks yang sangat besar (Jayasekara dan Abu, 2018). *Text mining* juga merupakan teknik analisis teks yang berasal dari dokumen. Tujuan dari teknik ini adalah untuk menemukan kata-kata yang dapat menggambarkan setiap arti yang ada di dalam dokumen sehingga dapat dilakukan analisis hubungan antar setiap kalimat. Teknik

dari *text mining* sendiri sudah banyak dikembangkan para peneliti karena teknik *text mining* sangat fleksibel digunakan untuk mengolah data teks (He dkk., 2013).

Area yang paling banyak digunakan untuk penambangan teks adalah:

1. Ekstraksi informasi (*information extraction*): Melibatkan pencocokan pola untuk menemukan hubungan dan frasa kunci dalam teks.
2. Pelacakan topik (*topic tracking*): Proses mengidentifikasi dokumen tambahan berdasarkan profil dan dokumen yang dilihat pengguna.
3. Perangkuman (*summarization*): Membuat ringkasan dokumen agar orang lebih mudah membacanya
4. Kategorisasi (*categorization*): Menentukan tema utama teks dan membagi teks ke dalam kategori tertentu berdasarkan temanya
5. Penggugusan (*clustering*): Kumpulan dokumen yang serupa tetapi tidak dikategorikan sebelumnya. (berbeda dengan kategorisasi di atas).
6. Penautan konsep (*concept linking*): Penautan dokumen yang berkaitan dengan identifikasi konsep yang dimiliki bersama membantu pengguna menemukan informasi yang mungkin tidak ditemukan dengan metode pencarian konvensional.
7. Penjawaban pertanyaan (*question answering*): Proses memberikan jawaban terbaik dengan mencocokkan pola berdasarkan pengetahuan.

Struktur data yang baik dapat memungkinkan komputerisasi secara otomatis. *Text mining* menggunakan informasi yang dapat distrukturkan dengan cara apa pun, jadi proses mengubah bentuk menjadi data yang terstruktur diperlukan untuk proses penggalian data, yang biasanya menghasilkan nilai numerik. Istilah "preprocessing teks" adalah yang paling sering digunakan untuk proses ini (Kadhim, 2018). Setelah terstruktur dan berbentuk nilai numerik, data dapat digunakan sebagai sumber untuk olahan lebih lanjut. Beberapa cara untuk menggambarkan proses *pre-processing text mining* adalah sebagai berikut: (Kadhim, 2018):

1. Semua huruf dalam dokumen diubah menjadi huruf kecil dengan metode *case folding*. Semua huruf dari "a" hingga "z" diterima. Karakter yang tidak termasuk huruf dihilangkan dan digunakan sebagai delimitter.

2. *Tokenizing*, merupakan tahap prosedur di mana suatu kalimat diuraikan berdasarkan kata yang menyusunnya.
3. *Filtering*, merupakan tahap dari proses pengambilan kata-kata kunci untuk hasil proses *token*. Anda dapat menggunakan algoritma untuk menyimpan kata penting/*wordlist* atau *stoplist* (membuang kata yang tidak penting). Kata-kata yang tidak memiliki deskripsi yang dapat dibuang dalam pendekatan *bag-of-words* disebut *stoplist* atau *stopword*.
4. *Stemming*, merupakan tahap di mana kata diuraikan menjadi bentuk kata dasarnya. Pada tahap ini, berbagai bentuk kata dikembalikan ke bentuk yang sama.
5. *Tagging*, merupakan tahap dalam proses mengubah kata lama menjadi kata dasar. Dalam bahasa Indonesia, proses ini tidak digunakan karena tidak mengenal kata lama.
6. *Analyzing*, merupakan tahap dari proses perhitungan bobot yang didasarkan pada tingkat kesesuaian kata kunci saat ini.

2.2.8 Extreme Programming

Extreme programming (XP) adalah salah satu metode pengembangan perangkat lunak dalam pendekatan *iterative process*. Agile development adalah kumpulan metode pengembangan perangkat lunak yang pengembangannya bersifat iteratif dan *incremental*. Agile development adalah metode pengembangan sistem yang digunakan (Pressman dan Maxim, 2015; Sommerville, 2016). *Extreme programming*, model pengembangan yang dikembangkan oleh Kent Beck, menggunakan pendekatan *agile* (Beck dan Andres, 2004). Perangkat lunak dikembangkan menggunakan pendekatan proses *agile* dengan desain yang sederhana, pengujian bertahap, dan dokumentasi yang minimal.

Planning, *design*, *coding*, dan *testing* adalah empat kerangka kegiatan dalam *extreme programming*, yang menggunakan pendekatan *object-oriented*. Kelebihan *extreme programming*:

1. Menjaga hubungan komunikasi yang baik dengan klien (*planning phase*)
2. Mengurangi pengeluaran pengembangan (*implementation phase*)

3. Meningkatkan interaksi dan rasa hormat antar developer (*implementation phase*)
4. XP adalah metodologi yang agak formal. Pembangun rencana harus selalu siap untuk perubahan karena perubahan akan selalu diterima, atau fleksibel. (*maintenance phase*)

Tidak dapat membuat kode yang rumit sejak awal, yang merupakan kelemahan dari prinsip kesederhanaan dan saran untuk melakukan apa yang diperlukan segera. Selain dari keunggulan dan kelemahan XP yang telah disebutkan di atas, XP juga memiliki keunggulan yang sekaligus menjadi kelemahan: selama pengembangan, tidak ada dokumentasi resmi yang dibuat. Satu-satunya instruksi yang tersedia adalah instruksi awal pengguna.

2.2.9 Uji Validitas dan Reliabilitas

Uji validitas menunjukkan seberapa valid data yang terjadi pada objek dibandingkan dengan sampel data yang dikumpulkan peneliti. Rumus yang digunakan untuk menghitung korelasi Pearson *Product Moment*, yang ditunjukkan dalam persamaan 2.1, digunakan untuk melakukan uji validitas. Uji validitas menggunakan korelasi Pearson *Product Moment*, yang merupakan korelasi antara item dan skor total dalam suatu variabel yang diukur. Tingkat signifikansi yang digunakan paling sering adalah 5% (Sugiyono, 2021, hal.176):

$$r_{xy} = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{\{(n\sum x^2) - (\sum x)^2\} \{(n\sum y^2) - (\sum y)^2\}}} \quad (2.1)$$

Dengan:

r_{xy} adalah nilai korelasi *Product Moment*

n adalah jumlah sampel

x adalah skor item

y adalah skor total

Perhitungan korelasi item-total yang benar (r hitung) harus lebih besar daripada r tabel (r hitung harus lebih besar daripada r tabel). Dalam penelitian ini, perhitungan ini dilakukan dengan menggunakan program SPSS. Dengan jumlah

sampel $n = 80$ dan taraf signifikansi sebesar 1%, diperoleh r tabel sebesar $= 0,286$ yang digunakan pada penelitian ini.

Sedangkan, uji reliabilitas menunjukkan seberapa konsisten dan stabil data suatu pengukuran. Saat digunakan untuk mengukur gejala yang sama, pengukuran yang reliabel akan menghasilkan hasil yang sebanding. Istilah "konsisten" digunakan untuk menggambarkan hal ini. Perhitungan nilai r untuk uji reliabilitas menggunakan metode *Cronbach Alpha*, menggunakan rumus yang sama seperti ditunjukkan dalam persamaan 2.2 (Sugiyono, 2021, hal.176):

$$r_{ac} = \frac{k}{(k-1)} \left\{ 1 - \frac{\sum \sigma_b^2}{\sigma_t^2} \right\} \quad (2.2)$$

Dengan:

r_{ac} adalah koefisien reliabilitas *Cronbach Alpha*

k adalah jumlah item pertanyaan yang diuji

$\sum \sigma_b^2$ adalah jumlah varian skor tiap item

σ_t^2 adalah varian total

Menurut Sugiyono, data dianggap reliabel jika koefisien reliabilitasnya lebih dari 0,6. Sumber lain mengatakan bahwa kategori reliabilitas yang lebih spesifik adalah kesepakatan umum yang dianggap cukup memuaskan jika nilai reliabilitasnya lebih dari 0,700. Jika nilainya kurang dari 0,700, item pertanyaan dianggap tidak reliabel. Tabel 2.1 menunjukkan klasifikasi kategori reliabilitas:

Tabel 2.1 Kategori Reliabilitas

Nilai <i>Alpha</i>	Keterangan
$Alpha > 0,90$	Reliabilitas Sempurna
0,70 – 0,90	Reliabilitas Tinggi
0,50 – 0,70	Reliabilitas Moderat (Cukup)
$Alpha < 0,50$	Reliabilitas Rendah (Tidak Reliabel)

2.2.10 Confusion Matrix

Confusion matrix adalah sebuah tabel yang digunakan untuk menggambarkan performa model klasifikasi pada set data uji yang nilai sebenarnya diketahui. *Confusion matrix* merupakan metode yang digunakan dalam pengujian akurasi pada suatu model sistem, yang pada dasarnya membandingkan hasil klasifikasi yang dihasilkan sistem model dengan hasil klasifikasi aktual. Terdapat 4 matriks dalam *confusion matrix* yang mengandung empat kombinasi nilai prediksi dan nilai aktual yang berbeda, yaitu nilai positif benar (TP), nilai positif salah (FP), nilai negatif salah (FN), dan nilai negatif benar (TN), seperti yang ditunjukkan dalam Tabel 2.2 di bawah ini (Azwar dkk., 2024).

Tabel 2.2 *Confusion Matrix*

		Predicted Class	
		Negatif	Positif
True Class	Negatif	True Negative (TN)	False Negative (FN)
	Positif	False Positive (FP)	True Positive (TP)

Penjelasan dari Tabel 2.2 model *confusion matrix* dijabarkan sebagai berikut:

- 1) *True positive* (TP), jumlah kasus di mana model memprediksi positif dan hasil sebenarnya juga positif
- 2) *True negative* (TN), jumlah kasus di mana model memprediksi negatif dan hasil sebenarnya juga negatif
- 3) *False positive* (FP), jumlah kasus di mana model memprediksi positif tetapi hasil sebenarnya negatif (disebut juga sebagai "*type I error*" atau "*false alarm*")
- 4) *False negative* (FN), jumlah kasus di mana model memprediksi negatif tetapi hasil sebenarnya positif (disebut juga sebagai "*type II error*" atau "*miss*")

Confusion matrix merupakan salah satu alat analitik prediktif yang menampilkan dan membandingkan nilai aktual atau nilai sebenarnya dengan nilai hasil prediksi model yang dapat digunakan untuk menghasilkan metrik evaluasi *accuracy*. Nilai akurasi sendiri adalah nilai yang menunjukkan keakuratan hasil

klasifikasi model sistem. Nilai ini digunakan untuk menghitung efisiensi sistem dan berguna untuk menentukan seberapa baik metode yang digunakan berhasil. Jadi, rasio prediksi yang benar (baik positif maupun negatif) terhadap semua data adalah ukuran akurasi. Berikut formula dari menghitung tingkat akurasi, pada persamaan 2.3:

$$Akurasi = \frac{(TP+TN)}{TP+TN+FP+FN} \times 100\% \quad (2.3)$$



SEKOLAH PASCASARJANA