

BAB II LANDASAN TEORI

2.1 Tinjauan Pustaka

Dalam penelitian ini memerlukan beberapa tinjauan pustaka yang diambil dari beberapa jurnal penelitian sebelumnya yang berkaitan dengan judul penelitian dan pokok pembahasan, yang nantinya digunakan sebagai abahan rujukan atau pendukung penelitian yang dilakukan, seperti yang tersaji berikut :

2.1.1 Tinjauan Terhadap Jurnal 01

Penelitian yang dilakukan oleh Fakhrurozi dan rekan-rekannya (2019) dari Fakultas Teknik Ilmu Komputer Universitas Teknokrat Indonesia membahas “Purwarupa aplikasi mesin penerjemah Bahasa Lampung dialek Api berbasis kamus dengan menggunakan pemrograman Python”. Pendekatan yang digunakan melibatkan pemecahan kalimat Bahasa Lampung menjadi daftar kata berdasarkan spasi, yang kemudian dicocokkan dengan kunci dalam database kamus. Hasil pencocokan ini menghasilkan nilai berupa kata dalam Bahasa Indonesia, yang selanjutnya disusun menjadi kalimat dalam Bahasa Indonesia. Tujuan utama penelitian ini adalah mengembangkan purwarupa aplikasi mesin penerjemah Bahasa Lampung dialek Api, sebagai solusi alternatif untuk menerjemahkan Bahasa Lampung dengan mempertahankan penggunaan kamus. Metode Waterfall digunakan dalam pengembangan aplikasi ini, dengan fokus utama pada algoritma penerjemahan kalimat. Aplikasi ini mampu menerjemahkan kalimat Bahasa Lampung sesuai dengan kosakata yang terdapat dalam kamus rujukan yang digunakan.

2.1.2 Tinjauan Terhadap Jurnal 02

Penelitian yang dilakukan oleh Rohman dan tim (2015) dari Jurusan Teknologi Informasi, Fakultas Teknik, Universitas Udayana membahas “Rancang Bangun Penerjemah Bahasa Indonesia ke Bahasa Jawa berbasis Android”. Dalam penjelasannya, penulis menyebut bahwa aplikasi penerjemah ini dikembangkan menggunakan metode rule-based dan brute force. Metode binary search digunakan sebagai media pencarian data, sementara algoritma Nazief-Andriani digunakan sebagai metode pengolahan kata dengan imbuhan. Aplikasi ini juga dilengkapi dengan penentuan unggahan Bahasa Jawa yang berfungsi sebagai penanda rasa hormat satu sama lain. Kemampuan aplikasi ini terletak pada kemampuannya menerjemahkan kata dan kalimat dari Bahasa Indonesia ke Bahasa Jawa pada tingkat tertentu sesuai dengan unggahan dasar yang ada. Waktu rata-rata penerjemahan per kata dalam proses tersebut adalah 152,94 milidetik, dengan tingkat akurasi sekitar 73,33%. Kualitas terjemahan yang dihasilkan juga menunjukkan angka yang tinggi, yakni sekitar 81,25%, di mana sebagian besar kata-kata pada input dapat diterjemahkan dengan benar.

2.1.3 Tinjauan Terhadap Jurnal 03

Penelitian yang dilakukan oleh Aryo Wibowo dan Mico Fahrizal (2021) berjudul "Metode Parsing Tree dalam Rancang Bangun Penerjemah Bahasa Indonesia ke Bahasa Lampung" bertujuan untuk mengembangkan aplikasi penerjemah bahasa Indonesia – Aksara Lampung sebagai upaya pelestarian dan pengenalan Aksara Lampung kepada masyarakat umum. Fokus penelitian ini adalah membangun aplikasi penerjemah yang memperkenalkan perbendaharaan kosakata dalam bahasa Indonesia dan bahasa Lampung beserta Aksara Lampung

kepada masyarakat. Metode yang digunakan dalam pembuatan aplikasi ini mencakup binary search dan parsing tree. Parsing tree digunakan untuk memecah kalimat menjadi kata-kata dan menggabungkan kata-kata menjadi kalimat, sementara binary search digunakan untuk mencari kosakata berdasarkan nilai indeks pada proses penerjemahan. Hasil penelitian menunjukkan bahwa aplikasi ini mampu menerjemahkan bahasa Indonesia – Aksara Lampung dalam bentuk kalimat tunggal yang diinputkan ke dalam aplikasi penerjemahan. Selain itu, aplikasi ini dapat menyajikan hasil pemecahan kalimat menjadi kata-kata beserta arti dari setiap kata tersebut. Kesimpulan dari penelitian ini adalah bahwa aplikasi penerjemah bahasa Indonesia – Aksara Lampung memiliki kemampuan untuk menerjemahkan kalimat tunggal dari bahasa Indonesia ke bahasa Lampung, lengkap dengan Aksara Lampung, serta berpotensi menjadi sarana untuk pemahaman dan pelestarian bahasa bagi masyarakat.

2.1.4 Tinjauan Terhadap Jurnal 04

Penelitian yang dilakukan oleh Karlana Indriani, dkk (2021) berjudul "Implementasi Penerjemah Bahasa Tetun – Indonesia (Timor Leste Tetun) Berbasis Web" bertujuan untuk mengatasi keterbatasan sumber daya penerjemah Bahasa Tetun, bahasa asal daerah pulau Timor, Timor Leste. Ketersediaan kamus atau aplikasi penerjemah untuk bahasa Tetun masih terbatas, sehingga menyulitkan individu yang ingin mempelajari bahasa ini atau hanya mencari informasi tentangnya. Oleh karena itu, penulis bertujuan untuk membantu menyediakan akses pembelajaran secara online yang dapat diakses secara fleksibel. Tujuan utama penelitian ini adalah mengembangkan aplikasi penerjemah bahasa Tetun – bahasa

Indonesia (Timor Leste Tetun) berbasis web. Penelitian ini menggunakan bahasa pemrograman PHP dan MySQL sebagai basis data penyimpanan kosakata bahasa Tetun dan bahasa Indonesia. Pendekatan yang diterapkan adalah Metode SDLC, yang melibatkan tahapan perencanaan sistem, tinjauan sistem, perancangan sistem, pengoperasian, dan pemeliharaan. Hasil penelitian menunjukkan bahwa aplikasi penerjemah bahasa Tetun – Indonesia ini memberikan kemudahan bagi pengguna untuk memperoleh pengetahuan tentang kosakata khususnya dalam bahasa Tetun – Indonesia.

2.1.5 Tinjauan Terhadap Jurnal 05

Penelitian yang dilakukan oleh Dian Dwi Ramadhan, dkk (2022) dengan judul “ Translator Bahasa Indonesia ke Bahasa Daerah (atau sebaliknya)”. Latar Belakang Penelitian ini dapat disimpulkan bahwa penggunaan mesin translator (terutama NMT) telah memberikan hasil yang positif dalam menerjemahkan bahasa Indonesia ke bahasa daerah atau sebaliknya. Mesin translator membantu mengatasi perbedaan bahasa antar daerah di Indonesia yang sangat beragam, sehingga masyarakat dapat lebih mudah memahami bahasa daerah tersebut. Namun, perlu diingat bahwa mesin translator juga memiliki keterbatasan, terutama dalam proses menangani nuansa dan konteks budaya tertentu yang mungkin sulit dipahami oleh mesin. Oleh karena itu, pengembangan mesin translator perlu terus dilakukan untuk meningkatkan akurasi dan kemampuan dalam menangani variasi bahasa daerah yang lebih kompleks. Studi ini juga menunjukkan bahwa terdapat variasi metode, baik menggunakan Statistical Machine Translation (SMT) maupun Neural Machine Translation (NMT), dengan masing-masing memiliki kelebihan dan kelemahan.

Oleh karena itu, pemilihan metode harus mempertimbangkan konteks dan kebutuhan spesifik dari penggunaannya. Sebagai langkah pengembangan ke depan, penelitian ini dapat menjadi dasar untuk eksplorasi lebih lanjut dalam meningkatkan kualitas mesin translator, termasuk penggunaan teknologi terbaru seperti machine learning dan deep learning. Selain itu, kerjasama antara ahli bahasa, pengembang perangkat lunak, dan peneliti teknologi dapat membantu menciptakan solusi yang lebih baik untuk kebutuhan penerjemahan bahasa Indonesia ke bahasa daerah atau sebaliknya.

2.1.6 Tinjauan Terhadap Jurnal 06

Penelitian yang dilakukan oleh Kadtabal (2016) dari Program Studi Teknik Informatika, STMIK Widya Cipta Dharma, membahas “Pembangunan aplikasi penerjemah kata (kamus) untuk Maluku Tenggara (Kei)–Indonesia berbasis Android”. Peneliti menjelaskan bahwa aplikasi ini bertujuan memberikan alternatif media pembelajaran dengan menggunakan ponsel pintar untuk memudahkan masyarakat dalam mencari kosa kata yang tepat. Metode yang digunakan dalam pembuatan aplikasi ini adalah metode pengembangan sistem waterfall. Tahapan ini meliputi perencanaan untuk menetapkan topik penelitian, analisis dan desain untuk mengumpulkan informasi dari berbagai sumber seperti buku dan artikel internet, serta perancangan aplikasi itu sendiri. Dengan persiapan yang matang, diharapkan aplikasi ini dapat memberikan kemudahan akses informasi dan meningkatkan pengetahuan pengguna.

2.1.7 Tinjauan Terhadap Jurnal 07

Penelitian yang dilakukan oleh Nurkholis, dkk (2019) dari Fakultas Teknologi Informasi, Universitas KH. A. Wahab Hasbullah berjudul "Aplikasi Kamus Bahasa Madura Berbasis Website". Dalam era teknologi informasi yang berkembang pesat, minat masyarakat terhadap pembelajaran bahasa Madura menurun, dan beberapa faktor yang mempengaruhi minat tersebut termasuk penurunan penggunaan bahasa Madura yang terlihat sebagai statis. Penelitian ini muncul dari keinginan untuk menciptakan aplikasi kamus bahasa Madura berbasis web guna memanfaatkan teknologi informasi dalam rangka melestarikan kebudayaan Indonesia. Tujuan utama penelitian ini adalah membangun aplikasi kamus bahasa Indonesia ke bahasa Madura dengan menggunakan platform berbasis web. Sistem yang dikembangkan difokuskan pada penyediaan fasilitas untuk menerjemahkan kata atau kalimat dari bahasa Indonesia ke bahasa Madura. Metode yang diterapkan dalam penelitian ini adalah algoritma pencarian sekuensial, yang merupakan pendekatan pencarian kata. Metode ini dipilih karena cocok untuk aplikasi kamus bahasa Indonesia- Madura, memungkinkan pencarian istilah dengan lebih cepat. Hasil dari penelitian ini adalah aplikasi kamus berbasis web yang dapat menerjemahkan kalimat atau kata dari bahasa Indonesia ke bahasa Madura. Aplikasi ini telah diuji dan membuktikan kemampuannya dalam melakukan penerjemahan dengan baik.

2.1.8 Tinjauan Terhadap Jurnal 08

Penelitian yang dilakukan oleh Yunusova Nilufar Maxmudjon Qizi (2023) dari Department of Foreign Languages of Namangan Institute of Engineering

Technology, Uzbekistan yang berjudul “Productivity of Machine Translation”. Penelitian ini membahas tentang bidang terjemahan teknis yang dikenal sebagai penerjemahan mesin (MT) atau penerjemahan berbantuan mesin. Metode terjemahan ini menggunakan berbagai jenis perangkat lunak komputer untuk menghasilkan terjemahan dari bahasa sumber ke bahasa target tanpa bantuan manusia. Terdapat berbagai metode penerjemahan mesin, dan banyak mesin penerjemah dalam bentuk mesin pencari gratis yang tersedia secara daring. Dalam bidang komunikasi teknis, terdapat dua jenis dasar penerjemah mesin yang mampu menerjemahkan jumlah teks besar sekaligus. Kedua jenis tersebut adalah penerjemah berbasis transfer dan penerjemah berbasis data. Sistem penerjemah berbasis transfer, yang mahal untuk dikembangkan, dibangun oleh ahli bahasa yang menentukan aturan tata bahasa untuk bahasa sumber dan bahasa target. Mesin bekerja dalam batasan aturan dan panduan yang dikembangkan oleh ahli bahasa.

2.1.9 Tinjauan Terhadap Jurnal 09

Penelitian yang dilakukan oleh Maisaroh (2019) dari Universitas Negeri Malang berjudul "Teknik Penerjemahan Kalimat Beridiom Dalam Kamus Saku Idiom Arab-Indonesia Karya Nuriyatul Hidayah" bertujuan untuk memberikan deskripsi mengenai teknik penerjemahan kalimat beridiom dalam kamus saku idiom bahasa Arab-Indonesia yang ditulis oleh Nuriyatul Hidayah. Penelitian ini menggunakan pendekatan kualitatif dengan fokus pada analisis pergeseran bentuk (transposisi) dan pergeseran makna (modulasi) pada kalimat beridiom. Data penelitian ini diperoleh dari kamus saku idiom bahasa Arab-Indonesia, dan hasilnya menunjukkan bahwa teknik penerjemahan kalimat beridiom dalam kamus tersebut

melibatkan pergeseran bentuk dan pergeseran makna. Metode pengumpulan data yang diterapkan dalam penelitian ini adalah teknik dokumentasi, di mana peneliti membaca dan mengamati teknik penerjemahan kalimat beridiom dalam kamus saku idiom Arab-Indonesia karya Nuriyatul Hidayah. Dengan demikian, penelitian ini memberikan kontribusi dalam memahami cara penerjemahan kalimat beridiom dalam kamus saku idiom bahasa Arab-Indonesia.

2.1.10 Tinjauan Terhadap Jurnal 10

Penelitian yang dilakukan oleh Muhammad Aziiz Pranaja, dkk (2023) dari Fakultas Vokasi, Universitas Negeri Surabaya yang berjudul “Pembuatan Aplikasi Mesin Penerjemah Menggunakan Metode No Language Left Behind Dari Bahasa Indonesia ke Bahasa Banjar”. Penelitian ini membahas pemanfaatan mesin penerjemah sebagai solusi bagi individu yang tertarik mempelajari bahasa di luar daerah asal mereka. Fokus utama penelitian adalah Bahasa Banjar, sebuah bahasa lokal yang, meskipun jarang digunakan secara aktif, memiliki signifikansi dalam literatur kuno, terutama di daerah Banjarmasin. Meskipun banyak mesin penerjemah yang telah dikembangkan, belum ada yang khusus untuk Bahasa Banjar. Dalam upaya mengatasi kekosongan ini, penelitian ini menyajikan BERT sebagai model yang sering digunakan dalam Natural Language Processing (NLP). Metode yang diimplementasikan adalah Sparsely Gated Mixture of Expert Layer, dan penulis berencana menggunakan berbagai library selama tahapan pembuatan aplikasi guna memudahkan proses penulisan. Pengujian sistem dilakukan dengan menggunakan matriks kebaikan BLEU. Selain itu, penulis juga mengukur waktu yang diperlukan untuk menerjemahkan kalimat atau teks dari Bahasa Indonesia ke

Bahasa Banjar. Hasil pengujian menunjukkan bahwa waktu yang dibutuhkan untuk satu kali menerjemahkan adalah sekitar 9,40 detik untuk prediksi dan 9,53 detik untuk mengakses API. Penelitian ini menciptakan kontribusi dalam memperluas pemahaman tentang penerjemahan bahasa, khususnya untuk Bahasa Banjar.

2.2 Landasan Teori

Terdapat beberapa teori dalam penelitian ini yang akan penulis jelaskan sebagai berikut:

2.2.1 Kamus

Menurut definisi dari Kamus Besar Bahasa Indonesia, Kamus adalah sebuah buku referensi yang berisi daftar kata dan ungkapan, biasanya disusun secara abjad, yang dilengkapi dengan keterangan mengenai makna, penggunaan, terjemahan, serta berisi kumpulan istilah atau nama yang disusun berdasarkan abjad dengan penjelasan tentang makna dan penggunaannya. Kamus berfungsi sebagai panduan untuk memahami arti kata atau ungkapan dan memberikan konteks penggunaannya. Asal usul kata "kamus" berasal dari bahasa Arab "*qāmus*," yang kemudian diadopsi dari bahasa Yunani "*okeanos*" yang berarti 'lautan'. Sementara itu, thesaurus adalah sebuah buku yang berisi daftar kata atau ungkapan yang memiliki hubungan makna. Istilah "*terasaurus*" diambil dari bahasa Yunani "*thērasauros*" yang berarti 'gudang'. Baik kamus maupun thesaurus diumpamakan seperti laut atau gudang yang berfungsi sebagai tempat penyimpanan atau penampungan untuk kekayaan kosakata dalam suatu bahasa. Keduanya mencerminkan 'kekayaan' dalam daya ungkap suatu bahasa (Badan Pengembangan dan Pembinaan Bahasa, 2016).

2.2.2 Bahasa Lampung

2.2.2.1 Pengertian Bahasa Lampung

Bahasa Lampung merupakan salah satu bahasa daerah di kepulauan nusantara yang masih dipelihara oleh para penutur bahasa yang dipakai oleh penduduk asli Lampung sebagai alat komunikasi antar masyarakat dalam kehidupan sehari-hari maupun acara adat istiadat. Bahasa Lampung terbagi dalam dua dialek yaitu dialek A (Api) dan dialek O (Nyo) (Septianingtias, 2012).

Provinsi Lampung, yang terletak di pulau Sumatra, memiliki posisi geografis yang sangat strategis. Dengan letaknya di ujung selatan Sumatera, provinsi ini menjadi pintu gerbang utama bagi mereka yang ingin mengakses Sumatra. Faktor ini menjadikan Lampung sebagai tujuan kunjungan yang menarik bagi pendatang dari berbagai wilayah dan latar belakang suku.

2.2.2.2 Dialek Bahasa Lampung

Menurut Agustina (2015), dalam bahasa Lampung di klasifikasikan rumpun bahasa dalam subdialek yaitu Dialek Belalau atau Dialek Api dan Dialek Abung atau Nyo.

1. Dialek Api

Dialek Api terbagi dalam Bahasa Lampung logat Belalau, Bahasa Lampung logat Krui, Bahasa Lampung logat Melinting, Bahasa Lampung logat Way Kanan, Bahasa Lampung logat Pubian, Bahasa Lampung logat Sungkai, Bahasa Lampung logat jelema daya atau logat Komerling.

2. Dialek Nyo

Dialek Nyo terbagi dalam Bahasa Lampung logat Abng dan Bahasa logat Menggala.

2.2.3 Karakteristik Bahasa Lampung

Bahasa Lampung mempunyai struktur tata bahasa yang serupa dengan Bahasa Indonesia. Di dalamnya terdapat subjek, predikat, objek, keterangan dan lain-lain. Kalimat dalam bahasa Lampung juga serupa dengan kalimat dalam Bahasa Indonesia, ada kalimat tunggal, kalimat majemuk, kalimat tanya, kalimat perintah, kalimat tanya, kalimat berita dan lain-lain (Megaria, 2013, dalam jurnal Jafar Fakhrurozi, Zaenal Abidin, Farida Ariyani). Kalimat bahasa Batak Lampung juga serupa dengan kalimat bahasa Indonesia, ada Kalimat Tunggal, Kalimat Majemuk, Kalimat Tanya, Kalimat Perintah, Kalimat Berita, Kalimat Sempurna dan lainnya. Struktur kalimat bahasa Lampung secara tulisan secara bentuk, umumnya terdiri dari:

Subjek (S) + Predikat (P) + Objek (O) + Kata Keterangan (K) = SPOK

Pada bagian ini disajikan berbagai kalimat dalam bahasa Lampung berikut terjemahannya dalam bahasa Indonesia :

a. Kalimat tunggal

Contoh dalam bahasa Lampung '*Adi lapah mit sekula*' yang artinya 'Adi pergi ke Sekolah'.

b. Kalimat Majemuk

Contoh dalam bahasa Batak '*Adi lapah mit sekulah walau badan ni mak sihat*' yang artinya 'Adi pergi ke Sekolah walau badan nya tidak sehat'.

c. Kalimat Perintah

Contoh dalam bahasa Batak '*Mengan Pai!*' yang artinya 'Makan dulu!'.

d. Kalimat Tanya

Contoh dalam bahasa Batak '*Ulah api niku miwang?*' yang artinya 'Karena apa kamu nangis?'.

e. Kalimat Berita

Contoh dalam bahasa Batak '*Ragah ino lagi maras janggukno*' yang artinya 'Lelaki itu lagi memapas janggutnya'

f. Kalimat Sempurna

Contoh dalam bahasa Batak '*Nyak haga ngebattu paman di nuwo*' yang arti nya 'Saya mau membantu paman di rumah'.

2.2.4 Penerjemahan

Setiap ahli memiliki interpretasi yang berbeda mengenai definisi penerjemahan. Beberapa definisi dari para ahli antara lain :

1. Daniels & Crystal (1992) menganggap penerjemahan sebagai istilah netral yang mencakup berbagai jenis tugas di mana makna ungkapan dalam satu bahasa diubah menjadi makna ungkapan dalam bahasa lain, baik melalui media lisan, tulisan, atau tanda. Mereka menekankan bahwa fokus utama dalam pengalihan bahasa adalah pada makna ungkapan, tanpa memandang media bahasa tertulis, lisan, atau isyarat.

2. Gutknecht & Gutknecht, yang dikutip oleh Aronof dan Miller, menyatakan bahwa penerjemahan umumnya merujuk pada materi tertulis, tetapi juga mencakup berbagai tugas di mana unsur-unsur teks dalam satu bahasa (bahasa sumber) diadaptasi ke dalam bahasa lain (bahasa sasaran), baik melalui medium tertulis, lisan, atau tanda.
3. Newmark, P. (1976) menyebutkan bahwa penerjemahan adalah keterampilan yang melibatkan usaha untuk mengganti pesan atau pernyataan tertulis dari satu bahasa ke bahasa lain, dengan tetap mempertahankan pesan atau pernyataan yang sama.
4. Catford (1965) mendefinisikan penerjemahan sebagai penggantian materi teks dari satu bahasa (bahasa sumber) dengan padanan materi teks dalam bahasa lain (bahasa sasaran).

Berdasarkan berbagai definisi ahli di atas, penerjemahan dapat diartikan sebagai suatu proses penulisan kembali atau ungkapan ulang pesan dari bahasa asal ke dalam bentuk bahasa tujuan atau sasaran tanpa mengurangi makna yang terkandung dalam bahasa asal.

2.2.5 *Database*

Database adalah suatu kumpulan data yang memiliki keterkaitan logis dan deskripsi, disimpan secara terstruktur dan dirancang untuk memenuhi kebutuhan informasi suatu organisasi. Prinsip utamanya terletak pada pengaturan data dan arsip, sementara tujuannya adalah untuk memberikan kemudahan dan kecepatan dalam pengambilan arsip, efisiensi ruang penyimpanan, mengurangi atau menghilangkan redudansi data, serta memastikan keakuratan informasi (Badriyah,

2008). Ada beberapa istilah umum yang sering dipakai pada *database*, yaitu sebagai berikut :

- a. *Field*, yaitu satu set kecil kata-kata atau deretan angka.
- b. *Record*, yaitu kumpulan dari *field* yang berelasi secara logis.
- c. *File*, yaitu kumpulan dari *record* yang berelasi secara logis.
- d. *Entity*, yaitu orang, tempat, benda atau kejadian yang berkaitan dengan informasi yang disimpan.
- e. *Attribute*, yaitu setiap karakteristik yang menjelaskan suatu *entity*.
- f. *Primary Key*, yaitu sebuah *field* yang nilainya unik yang tidak sama antara satu *record* dengan *record* yang lain.
- g. *Foreign Key*, yaitu sebuah *field* nilainya berguna untuk menghubungkan *Primary Key* yang berbeda pada table yang berbeda.

2.2.6 Bahasa Pemrograman *Python*

Python merupakan bahasa pemrograman tingkat tinggi yang diinterpretasikan, dirancang untuk pemrograman umum. Dikembangkan pertama kali oleh Guido van Rossum pada tahun 1991, setelah mengalami pengembangan sejak 1989. *Python* menggunakan pendekatan pemrosesan *interpreted*, di mana kode program dieksekusi baris per baris tanpa memerlukan proses kompilasi. Karakteristik ini serupa dengan bahasa skrip seperti PHP dan Javascript. *Python* dapat digunakan untuk berbagai tujuan pemrograman. *Python* banyak digunakan untuk membuat berbagai macam program seperti program CLI, Program GUI (Desktop), Aplikasi Mobile, Web, IoT, Game, Program untuk Hacking, dsb. *Python* juga dikenal dengan bahasapemrograman yang mudah dipelajari, karena struktur

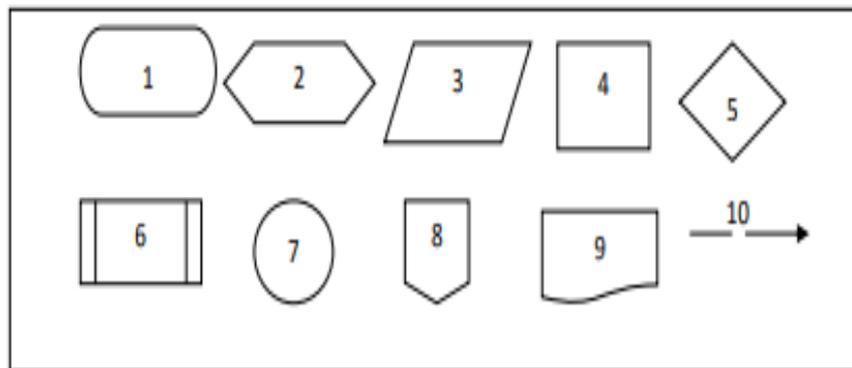
sintaksnya rapi dan mudah dipahami (Anwar, 2019).

2.2.7 Google Colaboratory

Google Colab atau *Google Colaboratory* pertama kali diperkenalkan oleh Google pada bulan April 2017. *Google Colab* dikembangkan untuk menyediakan lingkungan penelitian dan pengembangan yang mudah digunakan, terutama untuk pengembangan model Machine Learning dan Analisis Data, tanpa memerlukan konfigurasi yang rumit atau biaya infrastruktur yang tinggi. Platform ini memberikan akses ke sumber daya komputasi *Google Cloud* secara gratis yang memungkinkan pengguna untuk melakukan komputasi dan pelatihan model Machine Learning dengan cepat dan efisien.

2.2.8 Flowchart

Flowchart merupakan suatu alat yang digunakan untuk menggambarkan alur kerja penelitian, memetakan langkah-langkah dari perencanaan penelitian hingga analisis data (Creswell, 2008). *Flowchart* dapat diartikan diagram yang menggambarkan alur atau proses dalam menyelesaikan suatu tugas atau masalah secara grafis. Diagram ini menggunakan simbol-simbol standar untuk merepresentasikan berbagai jenis langkah atau tindakan dalam suatu proses, serta garis atau panah untuk menunjukkan urutan dan hubungan antar langkah tersebut. Bagan *flowchart* umumnya terdiri dari simbol-simbol *flowchart* seperti bentuk bangunan dua dimensi. Simbol nantinya akan dimasukkan ke dalam sebuah diagram yang detail dan ringkas. Simbol-simbol *flowchart* yang biasanya dipakai adalah simbol-simbol *flowchart* standar yang dikeluarkan oleh ANSI dan ISO. Simbol-simbol ini dapat dilihat pada gambar 2.1 berikut ini:



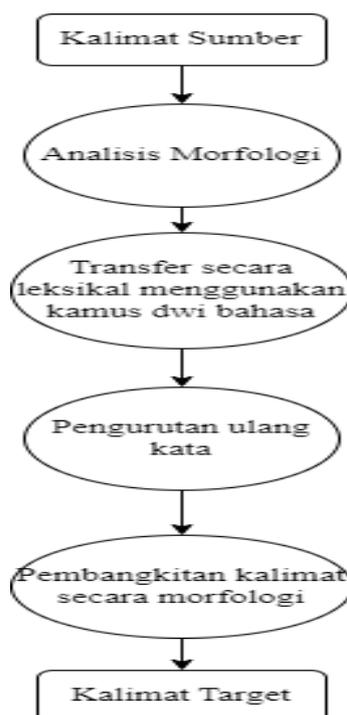
Gambar 2. 1 Simbol dalam Flowchart

Arti masing-masing symbol *flowchart* pada gambar 2.1 sebagai berikut :

1. Terminal, digunakan untuk memulai atau mengakhiri program.
2. Persiapan proses atau pemberian nilai awal
3. *Input* atau *Output*, digunakan untuk menyatakan masukkan dan keluaran tanpa melihat jenisnya.
4. Processing, digunakan untuk menunjukkan pengolahan yang akan dilakukan didalam komputer.
5. Decision, digunakan untuk memilih proses yang akan dilakukan berdasarkan kondisi tertentu.
6. Predefined Process, yaitu Rincian operasi berada di tempat lain.
7. Connector, digunakan untuk menyatakan sambungan dari proses yang satu ke proses berikutnya di halaman yang sama.
8. Offline Connector, Digunakan untuk menyatakan sambungan dari proses yang satu ke proses berikutnya di halaman yang berbeda.
9. Dokumen, yaitu *Input* atau *Output* yang dicetak.
10. Anak Panah, berfungsi untuk menghubungkan simbol yang satu dengan yang lainnya, menyatakan arus suatu proses.

2.2.9 *Direct Machine Translation*

Penerjemahan langsung melibatkan pengolahan kata per kata dari bahasa sumber ke bahasa tujuan dengan menggunakan kamus dwibahasa. Dalam proses ini, mesin penerjemah hanya melakukan *pra-pemrosesan* dan analisis morfologi yang dangkal tanpa memperhatikan struktur kalimat bahasa sumber. Kata-kata yang dihasilkan dari bahasa sumber kemudian dicocokkan satu per satu dengan kamus dwibahasa. Kata-kata yang memiliki padanan akan dikumpulkan dan disusun ulang sesuai tata susunan bahasa target. Langkah terakhir adalah menghasilkan terjemahan langsung dengan mempertimbangkan morfologi untuk mendapatkan kalimat yang sesuai dengan bahasa tujuan (Daniel Jurafsky, 2021).



Gambar 2.2 *Direct Machine Translation*

2.2.10 *BLEU SCORE*

BLEU (*Bilingual Evaluation Understudy*) merupakan algoritma yang berfungsi mengevaluasi kualitas dari sebuah mesin penerjemah yang telah

diterjemahkan oleh mesin dari satu bahasa alami ke bahasa lain. Poinnya ialah semakin dekat terjemahan mesin dengan terjemahan manusia maka akan semakin baik. Nilai *BLEU* dihasilkan dari perkalian *brevity penalty* dengan rata-rata geometri dari *modified precision score*. Semakin tinggi nilai *BLEU*, maka semakin akurat dengan rujukan. Nilai *BLEU* sendiri berada pada rentang 0 sampai 1. Suatu terjemahan akan mencapai nilai 1 jika terjemahan tersebut identik dengan terjemahan rujukan. Oleh karena itu, meskipun dengan penerjemahan oleh manusia tidak mungkin menghasilkan nilai 1. Sangat penting untuk diketahui bahwa semakin banyak terjemahan rujukan perkalamatnya, maka akan semakin tinggi nilainya. Untuk menghasilkan nilai *BLEU* yang tinggi, panjang kalimat hasil terjemahan harus mendekati panjang dari kalimat referensi dan kalimat hasil terjemahan harus memiliki kata dan urutan yang sama dengan kalimat referensi (Mulyana, Sujaini and Pratiwi, 2018). Berikut Rumus *BLEU* bisa dilihat pada gambar 2.3 yang terlampir.

$$\begin{aligned}
 BP_{BLEU} &= f(x) = \begin{cases} 1, & \text{if } c > r \\ e^{(1-r/c)}, & \text{if } c \leq r \end{cases} \\
 P_n &= \frac{\sum_{\text{coccopruen-gramoc}} \sum_{\text{count}_{allp}(n-gram)}}{\sum_{\text{coccopruen-gramoc}} \sum_{\text{count}(n-gram)}} \\
 BLEU &= BP_{BLEU} \cdot e^{\sum_{n=1}^N w_n \log P_n}
 \end{aligned}$$

Gambar 2. 3 Rumus *BLEU*

Keterangan :

BP : *brevity penalty*

c : jumlah kata dari hasil terjemahan otomatis

r : jumlah kata rujukan

Pn : *modified precision score*

Wn : $1/N$ (standar nilai N untuk *BLEU* adalah 4)

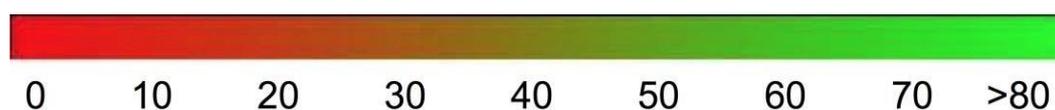
Pn : jumlah *n-gram* hasil terjemahan yang sesuai dengan rujukan dibagi jumlah *n-gram* hasil terjemahan.

BLEU Score yang telah didapatkan nilainya akan dinilai kualitas terjemahannya berdasarkan kategori nilai pada tabel 2.1 berikut :

Tabel 2.1 *BLUE Score* (<https://www.quora.com>)

| <i>BLEU Score</i> | Interpretasi |
|-------------------|--|
| <10 | Hampir tidak berguna |
| 10 – 19 | Sulit untuk mengerti intinya |
| 20 – 29 | Maksudnya dapat dimengerti namun susunan kalimatnya sangat kacau |
| 30 – 40 | Terjemahan yang dapat dimengerti |
| 40 - 50 | Terjemahan yang memiliki kualitas tinggi |
| 50 - 60 | Kualitas sangat tinggi, memadai, dan lancar |
| >60 | Melebihi terjemahan manusia |

Gambar kualitas rentang nilai dalam *BLEU*:



Gambar 2. 4 Rentang *BLEU Score* (Clouds, 2022)