

Model Pembelajaran Mendalam untuk Tugas NLP: Tinjauan Literatur Sistematis

Deep Learning Models for NLP Tasks: A Systematic Literature Review

Eduard Pangestu Wonohardjo¹⁾

¹ Computer Science, Binus Online Learning, Binus University

Jl. Kebon Jeruk Raya No. 27, Jakarta Barat 11530, DKI Jakarta

Telp : (021) 5345830

E-mail : eduard.wonohardjo@binus.ac.id¹⁾

Abstract

With the development of deep learning models, machine translation, which performs important tasks in natural language processing, has also made significant progress. Neural machine translation systems using deep neural network approaches have surpassed traditional statistical methods, especially with the emergence of transformer architectures and attention mechanisms. This systematic literature review examines recent developments in deep learning-based machine translation, focusing on multilingual pre-trained models, Transformers, BERT, GPT and Seq2Seq models from the last five years. Dealing with low-resource languages, training efficiency, and cross-domain translation quality are cited as key issues in neural machine translation. In this review article, we discuss its shortcomings. Additionally, this article also highlights how multilingual and unsupervised models have improved machine translation performance.

Keywords: *deep learning, machine translation, neural machine translation, seq2seq models, transformer models*

Abstrak

Dengan berkembangnya model pembelajaran mendalam, terjemahan mesin, yang melakukan tugas-tugas penting dalam pemrosesan bahasa alami, juga mengalami kemajuan yang signifikan. Sistem terjemahan mesin saraf yang menggunakan pendekatan jaringan saraf dalam telah melampaui metode statistik tradisional, terutama dengan munculnya arsitektur transformator dan mekanisme perhatian. Tinjauan literatur sistematis ini mengkaji perkembangan terkini dalam terjemahan mesin berbasis pembelajaran mendalam, dengan fokus pada model terlatih multibahasa, model Transformers, BERT, GPT, dan Seq2Seq selama lima tahun terakhir. Berurusan dengan bahasa yang miskin sumber daya, efisiensi pelatihan, dan kualitas terjemahan lintas domain disebut-sebut sebagai masalah utama dalam terjemahan mesin saraf. Dalam artikel ulasan ini, kami membahas kekurangannya. Selain itu, artikel ini juga menyoroti bagaimana model multibahasa dan tanpa pengawasan telah meningkatkan performa terjemahan mesin.

Kata kunci: *deep learning, machine translation, neural machine translation, seq2seq models, transformer models*

1. Pendahuluan

Terjemahan mesin (MT) selalu menjadi yang terdepan dalam aplikasi NLP, dengan tujuan memungkinkan komunikasi yang lancar antar bahasa. Awalnya, pendekatan berbasis aturan dan statistik seperti terjemahan mesin statistik (SMT) mendominasi MT. Sistem ini sangat bergantung pada corpora paralel dan fitur linguistik untuk menghasilkan terjemahan. Namun, dengan munculnya pembelajaran mendalam, bidang ini telah beralih ke model saraf yang lebih efisien, terukur, dan akurat, yang sering disebut sebagai Neural Machine Translation (NMT).

Pengenalan pembelajaran mendalam telah merevolusi MT dalam banyak hal, khususnya yang berkaitan dengan penggunaan jaringan saraf yang memungkinkan model mempelajari representasi bahasa dari data dalam jumlah besar. Sistem MT jaringan saraf awal didasarkan pada model sequence-to-sequence (Seq2Seq), yang mengkodekan kalimat sumber menjadi vektor dengan panjang tetap dan kemudian menerjemahkannya untuk menghasilkan kalimat target. Namun kelemahan model

ini adalah jika diterapkan pada kalimat yang panjang, terkadang informasi hilang selama proses pengkodean. Pengenalan mekanisme perhatian oleh Bahdanau et al. [1] mewakili terobosan signifikan yang memungkinkan model untuk fokus pada bagian berbeda dari kalimat masukan selama proses decoding.

Langkah transformatif selanjutnya datang dengan pengembangan model Transformer oleh Vaswani dkk. [2], yang sepenuhnya menggantikan jaringan saraf berulang dengan mekanisme perhatian otonom. Transformer telah menjadi tulang punggung sistem MT modern, menawarkan paralelisasi sistem yang lebih baik, waktu pelatihan yang lebih cepat, dan kualitas terjemahan yang sangat baik. Model bahasa terlatih seperti BERT dan GPT semakin meningkatkan kualitas terjemahan dengan memanfaatkan data tanpa pengawasan dalam jumlah besar, menjadikan sistem terjemahan lebih tangguh dan juga mudah beradaptasi antarbahasa.

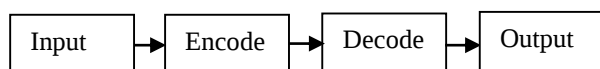
Karena pesatnya kemajuan teknologi terjemahan mesin akhir-akhir ini, makalah ini bertujuan untuk mengkaji

secara sistematis perkembangan terkini MT berbasis pembelajaran mendalam selama lima tahun terakhir. Perhatian khusus diberikan pada model seperti Transformers, BERT, GPT dan dampaknya terhadap peningkatan kualitas terjemahan dan penanganan bahasa dengan sumber daya terbatas. Selain itu, tinjauan ini membahas tren yang muncul seperti model multibahasa dan MT tanpa pengawasan yang telah melampaui batas-batas sistem penerjemahan tradisional. Dengan mengkaji kekuatan dan keterbatasan model-model ini, makalah ini bertujuan untuk memberikan pemahaman komprehensif tentang keadaan MT saat ini dan mengidentifikasi arah penelitian di masa depan.

2. Tinjauan Pustaka

2.1 Model Sequence-to-Sequence (Seq2Seq) dan Munculnya Mekanisme Perhatian

Pengenalan model sequence-to-sequence (Seq2Seq) menandai terobosan besar di bidang terjemahan mesin (MT). Metode tradisional seperti terjemahan mesin statistik (SMT) berbasis aturan dan sangat bergantung pada fitur buatan tangan dan corpora paralel. Sebaliknya, model Seq2Seq, pertama kali dikembangkan oleh Sutskever et al. [3] memperkenalkan kerangka end-to-end yang dapat secara langsung memetakan urutan masukan dalam bahasa sumber ke urutan keluaran dalam bahasa target. Model ini terdiri dari dua Recurrent Neural Networks (RNNs), satu bertindak sebagai encoder dan yang lainnya sebagai decoder, memungkinkan terjemahan dihasilkan tanpa aturan yang telah ditentukan sebelumnya (Gambar 2.1).



Gambar 1. Model Seq2Seq

Namun, keterbatasan utama model awal ini adalah ketidakmampuannya menangani ketergantungan jarak jauh, karena seluruh kumpulan sumber dikompresi menjadi vektor dengan panjang tetap oleh pembuat encode. Kemacetan ini mengakibatkan kualitas terjemahan yang lebih buruk untuk kalimat yang lebih panjang. Bahdanau dkk. [1] mengusulkan mekanisme perhatian sebagai solusi untuk masalah ini, yang memungkinkan model untuk secara dinamis memperhatikan bagian-bagian berbeda dari kalimat sumber selama proses decoding. Inovasi ini tidak hanya meningkatkan kualitas terjemahan, namun juga memungkinkan model Seq2Seq memproses rangkaian panjang dengan lebih efektif.

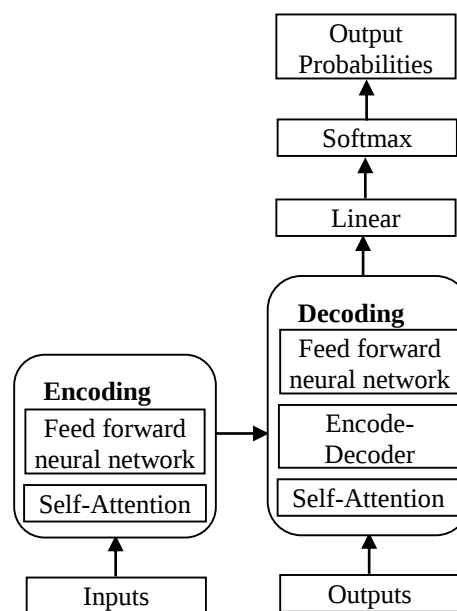
Mekanisme perhatian dengan cepat menjadi landasan terjemahan mesin saraf (NMT), yang menghasilkan peningkatan kinerja di berbagai pasangan bahasa. Bahdanau dkk. model yang diusulkan. sering disebut sebagai model Seq2Seq berbasis perhatian dan

kemampuannya untuk fokus pada bagian relevan dari rangkaian masukan telah menjadi bagian integral dari kemajuan modern di MT. Namun, meskipun sukses, model Seq2Seq yang mendapat perhatian masih mengalami inefisiensi komputasi karena ketergantungannya pada RNN.

Tobiyama dkk. [4] menjelaskan metode untuk mendeteksi infeksi di Windows dengan menilai ancaman suatu proses. Cara kerja metode ini adalah pertama-tama menangkap perilaku proses menggunakan Process Monitor dan kemudian membuat kumpulan fitur dari data ini menggunakan model Seq2Seq. Model Seq2Seq lainnya kemudian digunakan untuk mengklasifikasikan dan memperkirakan tingkat ancaman dari proses ini. Kemudian melakukan evaluasi menggunakan teknik validasi silang 5 kali lipat dan membandingkan efektivitasnya dengan metode sederhana menggunakan fungsi uni-gram.

2.2 Model Transformer: Pergeseran Paradigma dalam Terjemahan Mesin

Lompatan paling signifikan dalam MT terjadi dengan diperkenalkannya arsitektur Transformer oleh Vaswani dkk. [2], yang sepenuhnya menghilangkan kebutuhan akan pengulangan dalam jaringan saraf (Gambar 2). Transformer sepenuhnya didasarkan pada mekanisme perhatian diri, yang memungkinkan mereka memproses seluruh kalimat atau paragraf secara paralel, bukan secara berurutan, seperti halnya dengan RNN. Arsitektur ini telah menjadi dasar model terjemahan modern dan mendominasi bidang ini karena efisiensi komputasi dan kualitas terjemahannya yang unggul.



Gambar 2. Model Transformer

Mekanisme perhatian diri di Transformers memungkinkan model untuk mempertimbangkan arti

setiap kata dalam kalimat masukan relatif terhadap kata lain, sehingga memungkinkan pemahaman kontekstual yang lebih baik selama penerjemahan. Arsitektur Transformer juga menyertakan pengkodean posisi yang membantu model memahami urutan kata, yang merupakan faktor penting dalam terjemahan mesin. Sejak diluncurkan, Transformer tidak hanya meningkatkan kinerja terjemahan, tetapi juga mengurangi waktu pelatihan secara signifikan dibandingkan model berbasis RNN.

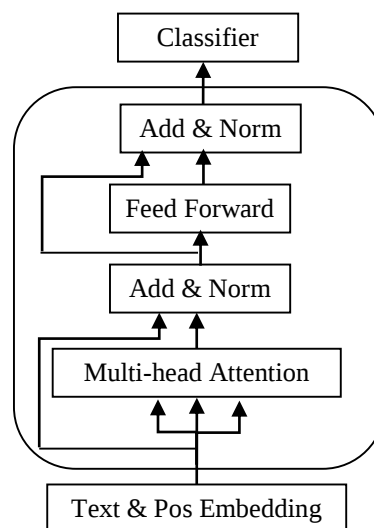
Dalam beberapa tahun terakhir, beberapa varian dan perbaikan pada Transformer telah diusulkan. Misalnya, Transformer-XL [5] memperluas arsitektur Transformer untuk menangani konteks yang lebih panjang dengan memperkenalkan pengulangan tingkat segmen, mengatasi batasan input dengan panjang tetap dalam Transformer standar. Liu dkk. [6] membahas X-Transformer, mekanisme perhatian diri yang merupakan bagian integral dari model Transformer, dan menekankan perannya dalam memahami struktur semantik kalimat dengan memberikan bobot perhatian yang berbeda pada setiap kata dibandingkan dengan kata lain. Mekanisme ini memungkinkan pemahaman konteks yang lebih baik, sehingga meningkatkan kualitas tugas seperti terjemahan mesin. Secara keseluruhan, kemampuan Transformers untuk fokus pada informasi yang relevan dalam data masukan mewakili kemajuan yang signifikan dalam aplikasi pemrosesan bahasa alami. Selain itu, meskipun BERT [7] terutama dirancang untuk tugas-tugas pemahaman bahasa, BERT telah diadaptasi untuk tugas-tugas penerjemahan karena mekanisme perhatian dua arah yang memungkinkan penyematan kontekstual yang kaya. A. Radford dkk. [8] telah membuktikan bahwa model bahasa dapat belajar melakukan tugas-tugas tertentu tanpa pengawasan eksplisit, terutama ketika dilatih pada kumpulan data yang besar, contohnya adalah jutaan halaman web. Hal ini menunjukkan efektivitas metode pelatihan tanpa pengawasan dalam pemrosesan bahasa alami. Dalam penelitian TB Brown dkk. [9] menyelidiki penerapan model bahasa berskala besar. Mereka menemukan bahwa penyempurnaan model ini secara signifikan meningkatkan kinerja dalam pembelajaran tanpa tugas dalam beberapa langkah, sehingga memungkinkan mereka menjadi lebih kompetitif dalam berbagai tugas dengan contoh yang minimal. GPT-2 dan GPT-3, yang didasarkan pada transformator searah, juga menjanjikan dalam penerjemahan mesin, khususnya dalam menghasilkan terjemahan yang lancar dan akurat secara kontekstual.

2.3 Model Bahasa terlatih dalam Terjemahan Mesin: BERT, GPT, dan Lainnya

Model bahasa terlatih seperti BERT, GPT, dan T5 telah memajukan bidang terjemahan mesin secara signifikan dengan memanfaatkan sejumlah besar data monolingual

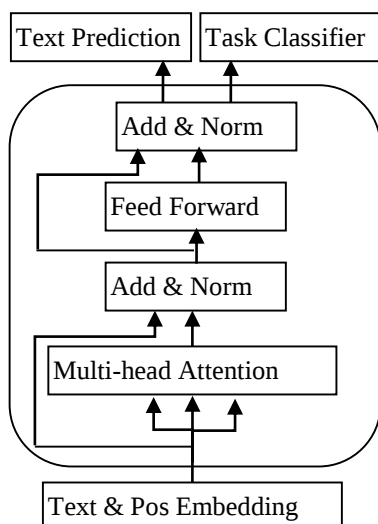
untuk mempelajari representasi linguistik yang kaya. Model ini, yang dilatih pada teks tak berlabel dalam jumlah besar, dapat disesuaikan untuk tugas tertentu, termasuk penerjemahan, menggunakan data berlabel dalam jumlah yang relatif kecil. Pendekatan ini sangat efektif dalam mengatasi masalah kelangkaan data dalam bahasa dengan sumber daya rendah.

BERT (*Bidirectional Encoder Representations from Transformers*) atau Representasi Encoder Dua Arah Transformers, diperkenalkan oleh Devlin dkk. [7] memiliki dampak besar pada berbagai tugas NLP, termasuk terjemahan mesin. Mekanisme perhatian dua arah BERT memungkinkannya menangkap konteks dari kedua arah (kiri dan kanan), yang sangat penting untuk memahami nuansa bahasa. Meskipun BERT tidak dirancang khusus untuk penerjemahan, BERT telah diadaptasi untuk tugas penerjemahan bersama dengan arsitektur dekoder tambahan, sehingga menghasilkan model yang berkinerja baik pada pasangan bahasa dengan sumber daya tinggi dan rendah (Gambar 3).



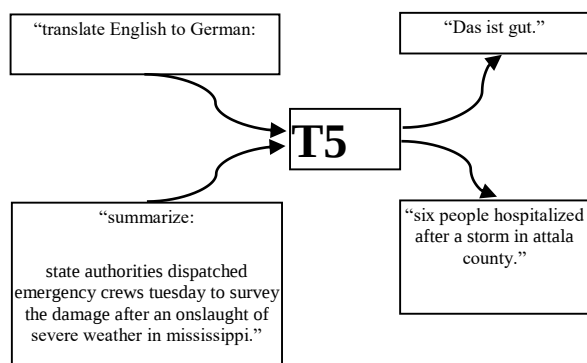
Gambar 3. Model BERT

Demikian pula, model GPT-2 dan GPT-3 (Generative Pre-Training Transformer) telah menunjukkan kemampuan pembuatan teks yang luar biasa, termasuk terjemahan (Gambar 4). Dilatih pada kumpulan data yang sangat besar, model ini dapat menghasilkan terjemahan yang lancar dengan akurasi tata bahasa yang tinggi. GPT-3 mewakili salah satu model bahasa terbesar dengan 175 miliar parameter [10] dan telah menunjukkan kinerja yang kuat pada tugas terjemahan zero-shot dan beberapa-shot menggunakan data pelatihan minimal.



Gambar 4. Model GPT

Model lain yang berpengaruh adalah T5 (Text-to-Text Transfer Transformer) oleh Raffel et al. [11], yang menyajikan setiap tugas NLP, termasuk terjemahan, sebagai masalah teks-ke-teks (Gambar 5). Kerangka kerja terpadu ini memungkinkan T5 unggul dalam penerjemahan mesin, terutama bila dikombinasikan dengan pra-pelatihan ekstensif tentang data multibahasa. L.Xue dkk. [12] memperkenalkan mT5, varian multibahasa dari model T5 (Text-to-Text Transfer Transformer). mT5 menggunakan kumpulan data baru yang berasal dari Common Crawl, mencakup 101 bahasa, dan dirancang untuk menangani semua tugas NLP berbasis teks menggunakan pendekatan teks-ke-teks terpadu. Arsitektur mT5 dibangun berdasarkan pendahulunya yang hanya satu bahasa dan memperluas kemampuannya ke berbagai bahasa lainnya.



Gambar 5. Model T5

2.4 Terjemahan Mesin Multibahasa dan Tanpa Pengawasan

Salah satu bidang penelitian MT yang paling menarik selama lima tahun terakhir adalah pengembangan model multibahasa yang mampu menerjemahkan berbagai bahasa tanpa pelatihan khusus tugas. Model seperti

mBERT (Multilingual BERT) dan XLM-R [13] telah menunjukkan kemampuan untuk mentransfer pengetahuan lintas bahasa, secara signifikan meningkatkan kualitas terjemahan untuk bahasa dengan sumber daya rendah. Model-model ini dilatih berdasarkan data dari berbagai bahasa, sehingga memungkinkan mereka mempelajari representasi lintas bahasa yang dapat diterapkan pada berbagai pasangan bahasa.

Selain model multibahasa, terjemahan mesin tanpa pengawasan telah muncul sebagai pendekatan yang menjanjikan untuk bahasa dengan corpora paralel terbatas. NMT tradisional sangat bergantung pada data paralel dalam jumlah besar untuk melatih model. Lample dkk. [14] memperkenalkan pendekatan tanpa pengawasan yang memanfaatkan corpora monolingual dan teknik seperti terjemahan balik untuk menghasilkan terjemahan tanpa memerlukan pasangan kalimat yang selaras. Ini sangat berguna untuk bahasa dengan sumber daya rendah di mana data paralelnya langka.

2.5 Terjemahan Mesin Dengan Sumber Daya Rendah: Tantangan dan Inovasi

Terjemahan mesin yang hemat sumber daya tetap menjadi salah satu bidang yang paling menuntut di MT. Meskipun bahasa yang kaya sumber daya seperti Inggris, Prancis, dan Tiongkok telah memperoleh manfaat besar dari model pembelajaran mendalam, bahasa dengan sumber daya yang lebih sedikit mengalami kesulitan untuk mencapai kinerja yang sebanding. Beberapa pendekatan inovatif telah diusulkan untuk memecahkan masalah ini.

Salah satu pendekatan tersebut adalah pembelajaran transfer, di mana model yang dilatih pada pasangan bahasa yang kaya sumber daya disesuaikan dengan pasangan bahasa yang miskin sumber daya, dengan memanfaatkan struktur bahasa yang umum. Terjemahan zero-shot, di mana model yang dilatih pada sekumpulan pasangan bahasa dapat menerjemahkan pasangan bahasa yang tidak terlihat tanpa pelatihan eksplisit, juga menunjukkan hasil yang menjanjikan. Metode lainnya adalah terjemahan balik, di mana data paralel sintetis dihasilkan dengan menerjemahkan data satu bahasa dari bahasa target ke bahasa sumber, sehingga menambah data pelatihan.

3. Metodologi Penelitian

Metodologi tinjauan literatur sistematis ini bertujuan untuk memastikan pemeriksaan komprehensif dan terstruktur terhadap kemajuan terkini dalam terjemahan mesin menggunakan model pembelajaran mendalam. Pendekatan ini mengikuti pedoman tinjauan sistematis yang umum, termasuk definisi pertanyaan penelitian, pemilihan sumber, ekstraksi data, dan analisis.

3.1 Pertanyaan Riset

Ulasan ini bertujuan untuk menjawab pertanyaan penelitian berikut:

1. Apa saja kemajuan terbaru dalam terjemahan mesin menggunakan model pembelajaran mendalam, khususnya yang berfokus pada Seq2Seq, Transformer, dan model terlatih seperti BERT dan GPT?
2. Bagaimana kinerja model terjemahan mesin multibahasa dan tanpa pengawasan dibandingkan dengan pendekatan yang diawasi?
3. Apa tantangan terbesar dan solusi inovatif ketika menerjemahkan bahasa yang kekurangan sumber daya?

3.2 Strategi Pencarian

Untuk memastikan cakupan komprehensif tentang kemajuan terkini dalam model pembelajaran mendalam untuk terjemahan mesin, pencarian sistematis dilakukan di beberapa database digital, termasuk:

- Google Scholar
- IEEE Xplore
- ACM Digital Library
- ScienceDirect
- arXiv (for preprints)

Kata kunci yang digunakan dalam penelusuran meliputi: “terjemahan mesin pembelajaran mendalam”, “model Seq2Seq”, “Terjemahan mesin transformator”, “terjemahan BERT”, “terjemahan mesin GPT”, “model terjemahan multibahasa”, dan “terjemahan mesin tanpa pengawasan”. Untuk membatasi cakupannya, pencarian dibatasi pada artikel yang diterbitkan antara tahun 2018 dan 2023 yang mencerminkan kemajuan terkini di bidang ini.

3.3 Kriteria Inklusi dan Eksklusi

Untuk memastikan relevansi dan kualitas sumber yang digunakan dalam tinjauan ini, kriteria inklusi dan eksklusi berikut diterapkan:

- Kriteria Inklusi:
 - Studi yang diterbitkan antara 2018 dan 2023.
 - Artikel yang berfokus pada model berbasis pembelajaran mendalam untuk terjemahan mesin.
 - Studi yang menyajikan hasil komparatif atau tolok ukur pada berbagai tugas terjemahan mesin.
 - Artikel yang mengeksplorasi penerapan model multibahasa, sumber daya rendah, atau tanpa pengawasan.
 - Artikel jurnal yang ditinjau sejawat, makalah konferensi, dan pracetak yang diterima secara luas.
- Kriteria Eksklusi:

- Artikel yang berfokus pada metode terjemahan mesin berbasis aturan atau statistik tanpa pembelajaran mendalam.
- Penelitian yang tidak mengandung hasil eksperimen atau empiris.
- Publikasi berbahasa non-Inggris, kecuali hasilnya relevan dengan tinjauan.

3.4 Ekstraksi dan Sintesis Data

Setiap studi terpilih ditinjau kontribusinya terhadap model pembelajaran mendalam dalam terjemahan mesin, dengan fokus pada aspek-aspek berikut:

- Arsitektur model (misalnya Seq2Seq, Transformer, BERT, GPT).
- Tugas-tugas yang menerapkan model tersebut (misalnya penerjemahan dengan sumber daya rendah, penerjemahan multibahasa, MT tanpa pengawasan).
- Hasil dan tolok ukurnya, terutama jika dibandingkan dengan metode yang paling canggih.
- Metode unik apa pun seperti terjemahan balik atau pembelajaran transfer yang digunakan untuk meningkatkan kinerja pada tugas tertentu.

Data tersebut kemudian disintesis untuk mengidentifikasi tema umum, tren, dan area untuk penelitian di masa depan. Analisis komparatif dilakukan untuk mengevaluasi kinerja model yang berbeda dalam kondisi berbeda, misalnya. B. antara bahasa yang kaya sumber daya dan miskin sumber daya serta pendekatan yang diawasi dan tidak diawasi.

3.3 Penilaian Mutu

Untuk memastikan keandalan tinjauan, setiap studi dinilai kualitas metodologisnya, termasuk:

- Kekokohan eksperimen dan hasil (misalnya ukuran kumpulan data, jumlah bahasa yang dievaluasi).
- Reprodusibilitas hasil (terlepas dari apakah kode sumber atau model telah tersedia untuk umum).
- Relevansi dan pentingnya hasil untuk bidang terjemahan mesin yang lebih luas.

4. Hasil dan Pembahasan

Pada bagian berisi penjelasan ilmiah dari hasil penerapan metode penelitian yang telah ditetapkan pada sub bab 3.

4.1 Temuan Utama

Tinjauan ini menyoroti beberapa pembelajaran penting di bidang terjemahan mesin menggunakan model pembelajaran mendalam, khususnya dalam lima tahun terakhir. Model berbasis transformator, khususnya

model Vaswani dkk. memperkenalkan Vanila Transformers. [2] telah menjadi arsitektur dominan dalam sistem MT modern, yang secara signifikan meningkatkan kualitas dan efisiensi terjemahan. Mekanisme perhatian diri, yang memungkinkan model untuk mempertimbangkan arti kata-kata yang berbeda dalam sebuah kalimat, telah terbukti menjadi kemajuan penting dalam menghasilkan terjemahan yang benar secara kontekstual. Transformer telah mengungguli model berulang tradisional seperti Seq2Seq, terutama ketika menangani ketergantungan jarak jauh.

Selain itu, munculnya model terlatih seperti BERT, GPT-2, GPT-3, dan T5 telah membuka peluang baru untuk penerjemahan mesin. Meskipun mekanisme perhatian dua arah BERT memberikan penyematan kontekstual yang kaya, kemampuan autoregresif GPT-3 telah menunjukkan hasil yang mengesankan dalam tugas penerjemahan zero-shot dan pagar-shot. Model-model ini tidak hanya meningkatkan kualitas terjemahan, namun juga mengurangi kebutuhan akan data paralel dalam jumlah besar, yang khususnya bermanfaat bagi bahasa-bahasa yang miskin sumber daya.

4.2 Perbandingan Performa Model Multibahasa dan Tanpa Pengawasan

Pengenalan model multibahasa seperti mBERT dan XLM-R telah secara signifikan meningkatkan penerjemahan ke berbagai bahasa, termasuk bahasa dengan sumber daya rendah. Model-model ini dilatih dalam berbagai bahasa secara bersamaan, memungkinkan mereka mentransfer pengetahuan dari bahasa yang kaya sumber daya ke bahasa yang miskin sumber daya, sehingga menghasilkan peningkatan kualitas terjemahan untuk bahasa yang miskin sumber daya. Model multibahasa seperti XLM-R [13] telah terbukti berkinerja baik di berbagai pasangan bahasa, sehingga mengurangi kesenjangan kinerja antara bahasa yang kaya sumber daya dan bahasa yang miskin sumber daya.

4.3 Tantangan dalam Terjemahan Mesin dengan Sumber Daya Rendah

Terlepas dari kemajuan ini, terjemahan mesin untuk bahasa dengan sumber daya rendah tetap menjadi tugas yang menantang. Meskipun model multibahasa dan pendekatan tanpa pengawasan telah mencapai kemajuan yang signifikan, model-model tersebut sering kali mengandalkan pembelajaran transfer dari bahasa-bahasa yang kaya sumber daya, yang mungkin tidak sepenuhnya menangkap nuansa bahasa-bahasa yang miskin sumber daya. Salah satu tantangan terbesarnya adalah kurangnya lembaga monolingual atau paralel yang besar dan berkualitas tinggi untuk melatih model-model ini. Selain itu, bahasa dengan sumber daya rendah mungkin memiliki struktur sintaksis atau kosakata unik yang tidak terwakili secara memadai dalam model bahasa sumber daya tinggi.

Terjemahan mesin tanpa pengawasan, yang menggunakan corpora monolingual dan teknik terjemahan balik, telah mendapatkan perhatian yang signifikan sebagai alternatif model yang diawasi yang memerlukan data paralel dalam jumlah besar. Lample dkk. [14] menunjukkan bahwa MT tanpa pengawasan dapat mencapai hasil kompetitif dengan metode yang diawasi pada pasangan bahasa tertentu. Namun, metode tanpa pengawasan cenderung mengalami kesulitan dengan struktur kalimat yang kompleks dan masih tertinggal dari model yang diawasi dalam hal kelancaran dan keakuratan tata bahasa.

Sennrich dkk. [15] menjelaskan bahwa metode inovatif seperti terjemahan balik dan pembelajaran transfer telah digunakan untuk menambah data pelatihan untuk bahasa yang miskin sumber daya. Namun, metode-metode ini hanya dapat menyelesaikan sebagian masalah kelangkaan data. Penelitian di masa depan diperlukan untuk mengeksplorasi cara-cara yang lebih efisien dalam mengumpulkan dan menghasilkan data berkualitas tinggi untuk bahasa-bahasa yang miskin sumber daya, misalnya melalui penggunaan data sintetis atau teknik crowdsourcing.

4.4 Arah Masa Depan

Masa depan penelitian terjemahan mesin terletak pada peningkatan efisiensi dan skalabilitas model pembelajaran mendalam, terutama untuk tugas penerjemahan multibahasa dan sumber daya rendah. Ada peningkatan minat untuk mengurangi biaya komputasi yang terkait dengan pelatihan model besar seperti GPT-3 dan meningkatkan kemampuan interpretasi dan penjelasan model ini. Selain itu, pertimbangan etis seputar penggunaan AI dalam penerjemahan, seperti: B. Bias dalam model bahasa dan potensi penyalahgunaan semakin mengemuka dan perlu dipertimbangkan dalam penelitian masa depan.

Rangkuman seluruh model beserta karakteristik, kekuatan dan kelemahannya dapat dilihat pada Tabel 1.

Tabel 1. Tabel Perbandingan Terjemahan Mesin Berbasis Deep Learning

Model	Karakteristik	Kekuatan	Kelemahan
Seq2Seq	Kerangka encoder-decoder; Mekanisme perhatian diperkenalkan untuk meningkatkan keselarasan dengan memfokuskan pada bagian-bagian yang relevan dari urutan masukan	Peningkatan akurasi terjemahan, terutama untuk urutan yang lebih panjang	Persyaratan sumber daya yang tinggi untuk pelatihan; Mekanisme perhatian meningkatkan kompleksitas komputasi

Model	Karakteristik	Kekuatan	Kelemahan
Transformer	Menggantikan RNN dengan mekanisme perhatian mandiri, memungkinkan pemrosesan urutan paralel	Kualitas dan efisiensi terjemahan yang lebih tinggi; Cocok untuk data dalam jumlah besar	Membutuhkan data pelatihan yang ekstensif; upaya komputasi yang tinggi
Pre-trained Language Model	Memanfaatkan pelatihan ekstensif sebelumnya; Fokusnya biasanya pada tugas bertopeng (BERT) atau autoregresif (GPT).	Pemahaman yang kuat tentang konteks; dapat beradaptasi dengan berbagai tugas NLP	Penyempurnaan diperlukan untuk kinerja spesifik MT; konteks searah dalam beberapa model
Multilingual and Unsupervised	Terlatih dalam berbagai bahasa atau tanpa corpora paralel; sering berbasis transformator	Efektif untuk bahasa dengan sumber daya rendah; Mengurangi kebutuhan corpora paralel	Mungkin kesulitan dengan pasangan bahasa yang berjauhan; Masalah transfer domain
Low-resource Machine Translation	Berfokus pada bahasa dengan data pelatihan terbatas; biasanya menggunakan transfer atau pembuatan data sintetis	Peningkatan akses ke MT untuk bahasa yang kurang terwakili; hemat biaya	Ketergantungan yang tinggi pada data sintesis; Kualitas bervariasi tergantung pada kelangkaan data

5. Kesimpulan

5.1 Simpulan

Terdapat kemajuan signifikan dalam penerjemahan mesin selama lima tahun terakhir, sebagian besar didorong oleh model *deep learning* seperti Transformer, BERT, dan GPT. Model-model ini tidak hanya meningkatkan kualitas terjemahan namun juga mencapai kemajuan yang signifikan dalam mengatasi tantangan penerjemahan yang terbatas sumber daya dan multibahasa.

5.2 Saran

Namun demikian, tantangan besar masih tetap ada, khususnya dalam bidang penerjemahan bahasa yang hemat sumber daya dan efisiensi komputasi model skala besar. Penelitian di masa depan kemungkinan besar akan berfokus pada optimalisasi model-model ini untuk

lingkungan yang miskin sumber daya, meningkatkan skalabilitasnya, dan mengatasi masalah etika yang terkait dengan penggunaannya.

Daftar Rujukan

- [1] D. Bahdanau, K. Cho, and Y. Bengio, "Neural Machine Translation by Jointly Learning to Align and Translate," CoRR, vol. abs/1409.0473, 2014, [Online]. Available: <https://api.semanticscholar.org/CorpusID:11212020>
- [2] A. Vaswani et al., "Attention is All you Need," in Neural Information Processing Systems, 2017. [Online]. Available: <https://api.semanticscholar.org/CorpusID:13756489>
- [3] I. Sutskever, O. Vinyals, and Q. Le, "Sequence to Sequence Learning with Neural Networks," Adv. Neural Inf. Process. Syst., vol. 4, 2014.
- [4] S. Tobiyama, Y. Yamaguchi, H. Hasegawa, H. Shimada, M. Akiyama, and T. Yagi, "A method for estimating process maliciousness with Seq2Seq model," in 2018 International Conference on Information Networking (ICOIN), 2018, pp. 255–260. doi: 10.1109/ICOIN.2018.8343120.
- [5] Z. Dai, Z. Yang, Y. Yang, J. Carbonell, Q. Le, and R. Salakhutdinov, "Transformer-XL: Attentive Language Models beyond a Fixed-Length Context," Oct. 2019, pp. 2978–2988. doi: 10.18653/v1/P19-1285.
- [6] H.-I. Liu and W.-L. Chen, "X-Transformer: A Machine Translation Model Enhanced by the Self-Attention Mechanism," Appl. Sci., vol. 12, no. 9, 2022, doi: 10.3390/app12094502.
- [7] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," Oct. 2018. doi: 10.48550/arXiv.1810.04805.
- [8] A. Radford et al., "Language models are unsupervised multitask learners," OpenAI blog, vol. 1, no. 8, p. 9, 2019.
- [9] T. B. Brown et al., "Language Models are Few-Shot Learners," ArXiv, vol. abs/2005.1, 2020, [Online]. Available: <https://api.semanticscholar.org/CorpusID:218971783>
- [10] OpenAI, "Pioneering Research on The Path to AGI." Accessed: Oct. 29, 2024. [Online]. Available: <https://openai.com/research/>
- [11] C. Raffel et al., "Exploring the limits of transfer learning with a unified text-to-text transformer," J. Mach. Learn. Res., vol. 21, 2020, [Online]. Available: <https://arxiv.org/abs/1910.10683>
- [12] L. Xue et al., "mT5: A massively multilingual pre-trained text-to-text transformer," 2021. [Online]. Available: <https://arxiv.org/abs/2010.11934>
- [13] A. Conneau, A. Baevski, R. Collobert, A. Mohamed, and M. Auli, "Unsupervised Cross-lingual Representation Learning for Speech Recognition," Oct. 2020. doi: 10.48550/arXiv.2006.13979.
- [14] G. Lample, M. Ott, A. Conneau, L. Denoyer, and M. Ranzato, "Phrase-Based & Neural Unsupervised Machine Translation," Oct. 2018, doi: 10.48550/arXiv.1804.07755.
- [15] R. Sennrich, B. Haddow, and A. Birch, "Improving Neural Machine Translation Models with Monolingual Data," Oct. 2015, doi: 10.48550/arXiv.1511.06709.