

BAB III

DESKRIPSI AI CHAT-GPT

A. Sejarah dan Perkembangan Chat-GPT

Chat-GPT merupakan bagian dari rangkaian model bahasa berbasis kecerdasan buatan yang dikembangkan oleh OpenAI, yang menggunakan teknologi pemrosesan bahasa alami (*Natural Language Processing/NLP*) untuk menghasilkan teks yang mirip dengan bahasa manusia. Pengembangan Chat-GPT dimulai dari upaya untuk membuat model yang mampu memahami dan merespons bahasa manusia secara lebih alami, melalui pelatihan dengan data teks besar dan arsitektur model yang kompleks, yaitu transformer.

1. GPT-1: Awal Perkembangan Model GPT

Model pertama dalam seri *Generative Pre-trained Transformer* (GPT), yang dikenal sebagai GPT-1, diperkenalkan oleh OpenAI pada tahun 2018. GPT-1 menggunakan pendekatan transformer untuk pertama kalinya, yang diperkenalkan dalam makalah "*Attention Is All You Need*" oleh Vaswani et al. (2017). Model ini dilatih dengan 117 juta parameter dari kumpulan data teks yang besar, yang memungkinkan GPT-1 untuk memahami pola bahasa secara lebih baik dibandingkan model sebelumnya. Meskipun GPT-1 telah menunjukkan kemampuan awal dalam menghasilkan teks yang koheren, keterbatasan dalam pemahaman konteks yang lebih luas masih terlihat pada tahap ini.¹

2. Chat-2 : Peningkatan Kapasitas dan Kemampuan Pemahaman Konteks

Pada tahun 2019, OpenAI meluncurkan GPT-2, model lanjutan yang dilatih dengan 1,5 miliar parameter, atau lebih dari sepuluh kali lipat dari GPT-1. Peningkatan kapasitas ini memungkinkan GPT-2 untuk menangkap konteks yang lebih dalam dalam teks, serta menghasilkan respons yang lebih panjang dan

¹ Ashish Vaswani dkk., "Attention Is All You Need" (arXiv, 2 Agustus 2023), 5998–6008, <https://doi.org/10.48550/arXiv.1706.03762>.

koheren. GPT-2 memperkenalkan kemampuan untuk menyelesaikan berbagai tugas tanpa pelatihan khusus, seperti menyelesaikan kalimat dan menjawab pertanyaan berdasarkan konteks. Kemampuan ini menunjukkan bahwa model transformer, bila ditingkatkan, dapat menangkap hubungan yang kompleks antar kata dan kalimat. Namun, GPT-2 masih memiliki keterbatasan, terutama dalam hal bias data dan ketidakmampuan untuk memahami konteks sosial atau budaya yang lebih mendalam.²

3. GPT-3 Ukuran dan Kapasitas

Diluncurkan pada tahun 2020, GPT-3 menjadi model bahasa terbesar yang dikembangkan oleh OpenAI pada saat itu, dengan 175 miliar parameter. Kapasitas yang sangat besar ini memungkinkan GPT-3 untuk menangani tugas-tugas yang lebih kompleks dan menghasilkan teks yang lebih menyerupai bahasa manusia dibandingkan model sebelumnya. GPT-3 dapat memahami dan merespons berbagai topik dengan lebih baik, dari konteks yang sederhana hingga yang kompleks. Selain itu, kemampuan GPT-3 dalam menghasilkan teks yang sesuai konteks telah memperluas aplikasinya di berbagai bidang, termasuk pendidikan, bisnis, dan layanan pelanggan. Meskipun demikian, GPT-3 masih menghadapi tantangan yang sama, yaitu keterbatasan dalam memahami konteks yang melibatkan nilai-nilai sosial atau norma yang lebih dalam.³

4. GPT-4 dan Versi Plus: Pemahaman Kontekstual dan Akurasi

Versi terbaru dari model ini, yaitu GPT-4 dan versi Plus, merupakan model dengan berbagai peningkatan dari GPT-3. GPT-4 menggunakan arsitektur yang lebih canggih dan data pelatihan yang lebih besar serta lebih beragam, yang memungkinkan untuk memberikan respons yang lebih akurat dan relevan terhadap konteks yang lebih kompleks. Salah satu keunggulan GPT-4 Plus adalah

² Alec Radford dkk., "Language Models are Unsupervised Multitask Learners," 2019, <https://www.semanticscholar.org/paper/Language-Models-are-Unsupervised-Multitask-Learners-Radford-Wu/9405cc0d6169988371b2755e573cc28650d14dfe>. 20

³ Tom B. Brown dkk., "Language Models are Few-Shot Learners" (arXiv, 2020), 1877–1901, <https://doi.org/10.48550/ARXIV.2005.14165>.

peningkatan dalam kemampuan pemahaman bahasa yang lebih halus, respons yang lebih cepat, dan efisiensi energi yang lebih baik. Selain itu, model ini dirancang untuk menghasilkan respons yang lebih aman dan etis dengan mengurangi risiko bias dan kesalahan interpretasi dalam konteks yang sensitif, termasuk dalam teks keagamaan atau materi berisiko lainnya.⁴

B. Teknologi Dasar: *Natural Language Processing* (NLP) dan Transformer

Teknologi utama yang mendasari Chat-GPT adalah *Natural Language Processing* (NLP) dan arsitektur model transformer. NLP merupakan cabang dari kecerdasan buatan yang berfokus pada kemampuan mesin untuk memahami, menghasilkan, dan berinteraksi menggunakan bahasa manusia. NLP bekerja dengan menguraikan teks menjadi unit-unit kecil, mengenali pola dalam data bahasa, dan menghasilkan respons yang sesuai berdasarkan pola tersebut.⁵ Dalam perkembangan Chat-GPT, NLP digunakan untuk memahami input pengguna, baik itu dalam bentuk perintah, pertanyaan, maupun konteks percakapan.

1. Model Transformer sebagai Inti dari NLP Modern

Arsitektur model transformer pertama kali diperkenalkan oleh Vaswani dan kawannya pada tahun 2017, dan model ini dengan cepat menjadi fondasi dari NLP modern⁶. Transformer menggunakan mekanisme *self-attention*, yang memungkinkan model untuk memberikan bobot berbeda pada setiap kata dalam suatu kalimat, berdasarkan konteks keseluruhan. Hal ini membantu model untuk menangkap makna yang lebih dalam dari teks, karena kata-kata di awal kalimat dapat memiliki hubungan dengan kata-kata di akhir. Dengan menggunakan *self-attention*, transformer memungkinkan Chat-GPT untuk memproses teks secara lebih efisien dan akurat dibandingkan dengan metode sebelumnya, seperti

⁴ Luciano Floridi dan Massimo Chiriatti, "GPT-3: Its Nature, Scope, Limits, and Consequences," *Minds and Machines* 30, no. 4 (Desember 2020): 681–94, <https://doi.org/10.1007/s11023-020-09548-1>.

⁵ Vaswani dkk., "Attention Is All You Need." 20

⁶ Radford dkk., "Language Models are Unsupervised Multitask Learners." 25

Recurrent Neural Networks (RNN) atau *Long Short-Term Memory* (LSTM) yang memiliki keterbatasan dalam memahami konteks yang panjang⁷.

2. Penerapan Transformer dalam Chat-GPT

Pada setiap interaksi model Chat-GPT, dari GPT-1 hingga GPT-4, OpenAI terus meningkatkan kapasitas dan kompleksitas model transformer yang digunakan. Model transformer yang digunakan oleh Chat-GPT dilatih dengan jutaan dokumen teks untuk mengenali pola bahasa alami dan memahami berbagai konteks secara lebih baik⁴. Dengan data pelatihan yang sangat besar, Chat-GPT dapat mengenali pola bahasa yang kompleks dan menghasilkan respons yang relevan dengan konteks input pengguna, termasuk dalam topik-topik rumit.

Namun, meskipun teknologi transformer memungkinkan Chat-GPT untuk menghasilkan teks yang menyerupai bahasa manusia, model ini tetap memiliki keterbatasan. Salah satunya adalah keterbatasan dalam memahami konteks historis dan nilai-nilai budaya, yang sangat penting dalam penafsiran teks keagamaan. Model transformer hanya dapat “memahami” konteks berdasarkan pola statistik, bukan melalui pemahaman yang mendalam seperti manusia⁸. Hal ini menyebabkan Chat-GPT seringkali hanya mampu memberikan respons yang umum, tanpa kedalaman interpretasi yang biasanya diperlukan dalam konteks spiritual atau keagamaan.

3. Peningkatan Pemahaman Kontekstual Melalui NLP

Chat-GPT 4o menggunakan arsitektur transformer terbaru, juga dilengkapi dengan berbagai algoritma NLP tambahan yang membantu meningkatkan akurasi dan relevansi dalam memahami input pengguna⁹. Model ini juga telah melalui proses pelatihan yang lebih ketat untuk meminimalkan bias dan kesalahan dalam menghasilkan respons. Namun, meski begitu, keterbatasan teknologi ini tetap ada, terutama ketika Chat-GPT harus memahami konteks-

⁷ Brown dkk., “Language Models are Few-Shot Learners.” 1877

⁸ Floridi dan Chiriatti, “GPT-3.” 681

⁹ “GPT-4,” diakses 10 November 2024, <https://openai.com/index/gpt-4/>.

konteks khusus yang memerlukan interpretasi nilai moral atau etika. Hal ini membuat teknologi NLP dalam Chat-GPT sangat terbatas ketika digunakan untuk menafsirkan teks-teks yang kompleks secara spiritual, seperti ayat-ayat Al-Qur'an.

C. Fitur dan Kemampuan Chat-GPT Secara Umum

Chat-GPT, sebagai model bahasa generatif yang dikembangkan oleh OpenAI, memiliki berbagai fitur dan kemampuan yang memungkinkan AI ini untuk memahami serta menghasilkan teks yang mirip dengan bahasa manusia. Chat-GPT tidak hanya digunakan untuk menjawab pertanyaan sederhana, tetapi juga dapat berfungsi sebagai asisten virtual, alat edukasi, dan bahkan sebagai pendukung dalam proses pengambilan keputusan dalam berbagai bidang.¹⁰

1. Kemampuan Pemahaman Bahasa

Salah satu fitur utama Chat-GPT adalah kemampuannya dalam memahami bahasa manusia dengan sangat baik. Berkat pelatihan dengan jutaan hingga miliaran parameter teks, Chat-GPT mampu menganalisis dan merespons teks berdasarkan konteks yang diberikan oleh pengguna. Kemampuan ini tidak hanya melibatkan pemahaman kata-kata secara individual, tetapi juga pemahaman yang lebih holistik terhadap keseluruhan kalimat dan paragraf. Model ini menggunakan mekanisme *self-attention* dalam arsitektur transformer untuk memperhitungkan konteks dari semua kata dalam kalimat, sehingga dapat memberikan respons yang lebih koheren dan relevan.¹¹

2. Kemampuan Menjawab Pertanyaan dan Menghasilkan Teks

Chat-GPT memiliki kemampuan luar biasa dalam menjawab berbagai pertanyaan, mulai dari pertanyaan faktual hingga pertanyaan yang membutuhkan penalaran. Kemampuan ini adalah hasil dari teknik *few-shot learning*, di mana Chat-GPT dapat menyelesaikan tugas baru meskipun hanya diberikan beberapa

¹⁰ Floridi dan Chiriatti, "GPT-3," 681–94.

¹¹ Vaswani dkk., "Attention Is All You Need," 5998–6008.

contoh atau bahkan tanpa contoh sama sekali¹². Selain itu, Chat-GPT mampu menghasilkan teks yang panjang, koheren, dan kreatif. Ini menjadikannya alat yang sangat berguna dalam pembuatan konten, penulisan artikel, dan pembuatan draf teks secara otomatis.¹³

3. Pemahaman dan Respons Kontekstual

Salah satu fitur yang membedakan Chat-GPT dengan model bahasa lainnya adalah kemampuannya untuk mempertahankan konteks percakapan. Dalam interaksi multi-turn (beberapa putaran percakapan), Chat-GPT dapat merespons dengan mempertimbangkan pesan-pesan sebelumnya, sehingga percakapan menjadi lebih alami dan respons yang diberikan tidak terputus atau terlepas dari konteks yang diinginkan pengguna.¹⁴ Kemampuan ini membuat Chat-GPT cocok untuk digunakan dalam aplikasi chatbot atau layanan pelanggan otomatis.

4. Fleksibilitas dalam Menyelesaikan Berbagai Tugas

Chat-GPT adalah model yang fleksibel dan mampu menjalankan berbagai jenis tugas dalam satu model. Misalnya, selain menjawab pertanyaan dan membuat teks, Chat-GPT juga bisa melakukan penerjemahan bahasa, penyelesaian kode, serta penulisan esai atau bahkan analisis sentimen dalam teks. Fleksibilitas ini menjadikannya alat yang berguna untuk beragam aplikasi, mulai dari pendidikan hingga bisnis.¹⁵

5. Keterbatasan dalam Pemahaman Nilai Sosial dan Budaya

Meskipun Chat-GPT memiliki kemampuan luar biasa dalam pemahaman bahasa dan respons kontekstual, model ini tetap memiliki keterbatasan dalam

¹² Brown dkk., "Language Models are Few-Shot Learners," 1877–1901.

¹³ Radford dkk., "Language Models are Unsupervised Multitask Learners." 25

¹⁴ Yupeng Chang dkk., "A Survey on Evaluation of Large Language Models" (arXiv, 28 Agustus 2023), 19, <http://arxiv.org/abs/2307.03109>.

¹⁵ Emily M. Bender dan Alexander Koller, "Climbing towards NLU: On Meaning, Form, and Understanding in the Age of Data," dalam *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, Online: Association for Computational Linguistics, 2020), 5185–98, <https://doi.org/10.18653/v1/2020.acl-main.463>.

menangani konteks sosial, budaya, dan nilai-nilai yang kompleks. Chat-GPT dilatih menggunakan data teks yang sangat besar dari internet, yang berarti ia hanya mengenali pola-pola dalam teks tanpa benar-benar memahami nilai atau makna mendalam yang ada dalam budaya tertentu.¹⁶ Ini menyebabkan Chat-GPT kadang-kadang menghasilkan respons yang kurang sesuai atau tidak sensitif terhadap konteks sosial atau budaya tertentu. Misalnya, dalam menafsirkan teks agama, Chat-GPT mungkin memberikan respons berdasarkan pola umum tanpa pemahaman nilai spiritual atau historis yang diperlukan.¹⁷

D. Kemampuan Chat-GPT 4o

Chat-GPT 4o merupakan versi yang telah mengalami peningkatan signifikan dibandingkan model sebelumnya. Dengan beberapa pembaruan yang dirancang untuk meningkatkan kapasitas pemahaman, kecepatan respons, dan akurasi, Chat-GPT versi Plus menjadi lebih unggul dan relevan untuk aplikasi yang lebih kompleks dan kritis. Pembaruan pada versi Plus ini memberikan peningkatan pada beberapa aspek utama yang penting, terutama dalam penerapan AI untuk tugas-tugas yang memerlukan respons lebih kontekstual dan bernuansa.

1. Perbedaan dengan Versi 3.5

Salah satu perbedaan utama antara Chat-GPT versi 4o dengan versi sebelumnya adalah pada jumlah parameter dan data pelatihan yang lebih besar. Chat-GPT versi Plus dilatih menggunakan data yang lebih banyak dan lebih beragam, sehingga mampu menangkap nuansa bahasa dan konteks dengan lebih baik¹⁸. Peningkatan jumlah parameter ini memungkinkan versi Plus untuk memahami hubungan yang lebih kompleks antara kata-kata dan frasa dalam teks, yang memberikan respons yang lebih akurat dan mendalam dibandingkan dengan versi sebelumnya.¹⁹

¹⁶ Bender dan Koller.

¹⁷ Avanti Shrikumar, Peyton Greenside, dan Anshul Kundaje, "Learning Important Features Through Propagating Activation Differences," 2017, <https://doi.org/10.48550/ARXIV.1704.02685>.

¹⁸ Brown dkk., "Language Models are Few-Shot Learners," 1877–1901.

¹⁹ Radford dkk., "Language Models are Unsupervised Multitask Learners." 26

2. Kemampuan Pemahaman Bahasa yang Lebih Mendalam

Chat-GPT versi 4o memiliki kemampuan untuk memahami bahasa dengan lebih halus dan mendalam, terutama dalam konteks yang lebih spesifik dan bernuansa. Model ini telah disesuaikan untuk memahami frasa ambigu dan konteks percakapan yang lebih panjang, membuatnya lebih efektif dalam percakapan multi-turn di mana konteks percakapan perlu dipertahankan.²⁰ Peningkatan ini juga didukung oleh algoritma tambahan yang mampu mengidentifikasi konteks yang lebih baik dan respons yang lebih relevan dalam percakapan yang kompleks.²¹

3. Kecepatan dan Efisiensi Respons yang Lebih Baik

Selain pemahaman bahasa yang lebih baik, Chat-GPT versi Plus juga memiliki peningkatan dalam hal kecepatan dan efisiensi. Versi Plus ini dirancang untuk memberikan respons lebih cepat dibandingkan versi sebelumnya, terutama ketika digunakan dalam aplikasi interaktif seperti chatbot atau asisten virtual.²² Peningkatan efisiensi ini dicapai dengan optimalisasi algoritma pemrosesan, yang memungkinkan model untuk menghasilkan respons dalam waktu yang lebih singkat, sehingga lebih nyaman dan efisien bagi pengguna.²³

4. Pengurangan Bias dan Keamanan yang Lebih Tinggi

Salah satu keunggulan utama dari Chat-GPT versi Plus adalah peningkatan pada aspek keamanan dan pengurangan bias. Versi ini dirancang untuk menghasilkan respons yang lebih aman dan etis, dengan lebih sedikit risiko bias yang tidak diinginkan atau respons yang tidak sesuai konteks. OpenAI telah menerapkan filter tambahan dan metode pelatihan yang lebih terkontrol untuk meminimalkan kemungkinan respons yang bermuatan bias sosial atau budaya. Hal

²⁰ Floridi dan Chiriatti, "GPT-3," 681–694.

²¹ "GPT-4."

²² Vaswani dkk., "Attention Is All You Need," 5998–6008.

²³ Bender dan Koller, "Climbing towards NLU." 5185-5198

ini sangat penting untuk aplikasi yang digunakan dalam konteks yang sensitif, seperti penafsiran teks keagamaan atau dalam interaksi publik yang luas.²⁴

5. Aplikasi yang Lebih Luas dan Kemampuan Adaptasi yang Tinggi

Dengan kemampuan adaptasi yang lebih baik, Chat-GPT versi Plus cocok untuk aplikasi yang lebih kompleks, seperti pendidikan, analisis data, dan dukungan keputusan. Versi ini dapat disesuaikan untuk memenuhi kebutuhan pengguna dalam berbagai industri, mulai dari perbankan, layanan kesehatan, hingga pendidikan, di mana pemahaman konteks dan respons yang akurat sangat dibutuhkan²⁵. Keunggulan ini menjadikan Chat-GPT versi Plus sebagai pilihan yang lebih andal untuk aplikasi yang memerlukan interaksi bernuansa dan responsif terhadap konten.

²⁴ Shrikumar, Greenside, dan Kundaje, "Learning Important Features Through Propagating Activation Differences." 20

²⁵ "GPT-4."

