

## 2 LANDASAN TEORI

### 2.1 Tinjauan Pustaka

Berikut berisikan tinjauan pustaka yang digunakan dalam pelaksanaan penelitian ini

#### 2.1.1 Studi Alkitab

Studi Alkitab adalah aplikasi akademik serangkaian disiplin ilmu untuk mempelajari Alkitab. Ada beberapa ayat Alkitab yang mendasari pentingnya untuk melakukan studi Alkitab, “Semua Kitab Suci dinapasi oleh Allah dan bermanfaat untuk mengajar, untuk menyatakan kesalahan, untuk memperbaiki kelakuan, dan untuk mendidik dalam kebenaran” (2 Timotius 3:16.AYT), “Tetapi Yesus menjawab : "Ada tertulis: Manusia hidup bukan dari roti saja, tetapi dari setiap firman yang keluar dari mulut Allah” (Matius 4:4.AYT). Studi Alkitab merupakan salah satu cara untuk kita dapat mendapatkan maksud dari firman-Nya dan menerapkannya dalam hidup kita. Maka dari itu penting bagi orang percaya untuk terus menerus melakukan studi Alkitab.

Ada beberapa cara dalam melakukan studi Alkitab, antara lain sebagai berikut:

a. Studi Alkitab tradisional

Studi Alkitab tradisional dilakukan menggunakan Alkitab fisik, pena dan kertas. Studi Alkitab tradisional relatif susah untuk dilakukan dikarenakan harus membawa Alkitab fisik.

b. Studi Alkitab digital

Studi Alkitab digital dilakukan dengan menggunakan Alkitab dalam bentuk aplikasi, dan segala bahan yang sudah digital. Studi Alkitab digital relatif mudah untuk dilakukan dikarenakan hanya perlu membawa *handphone*. Studi Alkitab digital juga kini dapat dilakukan menggunakan *Generative AI* seperti Google Bard, ChatGPT, Perplexity AI dan lain lain.

c. Studi Alkitab Kontekstual

Studi Alkitab Kontekstual adalah metode studi Alkitab interaktif yang melibatkan dialog antara konteks pembaca dengan konteks Alkitab. Studi Alkitab Kontekstual dikembangkan di Afrika Selatan pada 1980an sebagai "hermeneutika pembebasan". Metode ini memakai pendekatan "Lihat, Pertimbangkan, Bertindak" yang terdiri dari

menganalisis konteks, merenungkan teks Alkitab, dan menindaklanjuti hasil renungan tersebut. Tujuan dari Studi Alkitab Kontekstual adalah untuk mencapai transformasi pribadi dan komunal. Salah satu contoh penerapannya adalah dengan melibatkan jemaat untuk merenungkan isu lokal melalui diskusi dipimpin mengenai teks Alkitab yang terkait.

d. Studi Alkitab Antarbudaya

Studi Alkitab Antarbudaya merupakan variasi dari Studi Alkitab Kontekstual yang melibatkan kelompok-kelompok dari latar belakang budaya dan sosial yang berbeda untuk membahas teks Alkitab bersama-sama. Pendekatan ini memungkinkan saling belajar dari beragam sudut pandang dan pengalaman yang dimiliki masing-masing kelompok. Salah satu contoh penerapannya adalah dengan mengumpulkan kelompok Kristen dari berbagai negara untuk melakukan pembahasan daring mengenai topik Alkitab yang sama.

### 2.1.2 Chatbot

*Chatbot* adalah salah satu contoh dari produk komunikasi antar manusia dan komputer (Bansal, H. & Khan, R; 2018). *Chatbot* merupakan salah satu produk dari *Artificial Intelligence* yang dapat memungkinkan manusia untuk berbicara dengan komputer. *Chatbot* yang ditemukan pada umumnya terdiri dari 3 jenis yaitu *chatbot rule-based*, *chatbot machine learning* dan *chatbot LLM*.

#### 2.1.2.1 Chatbot rule based

*Chatbot rule-based* adalah jenis *chatbot* yang menggunakan serangkaian aturan yang telah ditentukan sebagai dasar untuk menjawab suatu pertanyaan (Adamopoulou, E. & Moussiades, L; 2020). *Chatbot* jenis ini adalah *chatbot* yang kaku karena *chatbot* ini tidak dapat menjawab pertanyaan di luar aturan yang telah ditentukan. *Chatbot* jenis ini juga dikenal sebagai *decision-tree bot* atau *linguistic-based chatbot*. *Chatbot rule-based* cocok untuk perusahaan yang memiliki tujuan tertentu, dan tidak memerlukan banyak percakapan.

#### 2.1.2.2 Chatbot machine learning

*Chatbot machine learning* adalah sebuah *chatbot* dirancang untuk meniru percakapan manusia. *Chatbot* ini menggunakan teknologi *machine learning* untuk menganalisis data dan belajar sendiri untuk berinteraksi seperti manusia. *Chatbot* ini dapat memproses pertanyaan

dan memberikan respon sebagai jawaban atas pertanyaan tersebut. Chatbot ini dapat membantu dalam pekerjaan yang membutuhkan kemampuan berkomunikasi dengan manusia. *Chatbot* ini dapat *ditraining* agar bahasanya tidak kaku.

### 2.1.3 Open AI

OpenAI adalah sebuah perusahaan yang bergerak di bidang pengembangan *Artificial Intelligence* (AI). OpenAI memiliki banyak produk seperti ChatGPT, DALL-E, Whisper, TTS dan masih banyak lagi. OpenAI juga menyediakan API yang dapat memungkinkan programmer untuk memakai produk - produk OpenAI didalam aplikasinya. Selain mengembangkan *Custom GPT*, OpenAI juga menyediakan *Assistant* API yang memungkinkan programmer untuk menghubungkan data - data eksternal dengan ChatGPT (OpenAI, 2023). Sehingga ChatGPT dapat menjadi alat yang *powerful* dikarenakan dapat dihubungkan secara langsung dengan data - data eksternal.

#### 2.1.3.1 ChatGPT/Chat Completion

Sebuah model chatbot yang bisa *generate text* berdasarkan *prompt* yang kita berikan. *Inputnya* berupa JSON yang berisikan parameter:

- *model* (ChatGPT 3.5 *Turbo*, ChatGPT 4, ChatGPT 4 *Turbo*, ChatGPT 4 *Turbo Vision*)
- *message* (*list message* yang berisikan *role* (*user*, *system*, *assistant*, *function*) dan *message*)
- *max\_tokens* (*max token* dari *response* ChatGPT)
- *frequency\_penalty* (dari -2.0 sampai 2.0, untuk mengatur jumlah pengulangan kata)
- *n* (jumlah respon chat)
- *response\_format* (tersedia *json\_object* supaya hasil *responsenya* dalam *json objects* bukan dalam bentuk *text*)
- *stream* (jika *true* maka akan menghasilkan response per suku kata, jika *false* maka akan menghasilkan *response* yang utuh)
- *top\_p* (0 sampai 1, mempengaruhi seberapa banyak kemungkinan yang dipertimbangkan model saat memilih token berikutnya)
- *seed* (jika *parameter* dan *seed* sama maka *response* yang dikeluarkan untuk 1 *prompt* yang sama akan sama)
- *function\_call* (untuk memasukkan *function* - *function* yang dapat digunakan oleh *chat*)

*Output* yang didapatkan berupa JSON yang berisikan *parameter*:

- *id*
- *object*
- *model*
- *choices* (yang berisikan *message* (*role*, *context*) dan *finish\_reason*)
- *usage* (yang berisikan *prompt\_tokens*, *completion\_tokens* dan *total\_tokens*)

### 2.1.3.2 Assistant API

*Assistant* merupakan salah satu API OpenAI yang terbaru dan masih *beta*. *Assistant* menggunakan *chat model* namun kita tidak perlu menyimpan *history message* karena *history* sudah disimpan oleh *Assistant* di dalam *thread*. *Assistant* juga memiliki beberapa keunggulan yaitu mempunyai *data retrieval* (mengambil data eksternal dalam bentuk pdf, csv, json dan masih banyak lagi) dan *multiple function calling* yang memungkinkan *assistant* untuk menggunakan fungsi dalam program kita. Untuk membuat *assistant* baru membutuhkan JSON dengan *parameter* sebagai berikut:

- *model* (pilih salah satu dari *chat completion model*)
- *name* (nama dari *assistant* yang akan dibuat, maksimal 256 karakter)
- *description* (deskripsi dari *assistant*, maksimal 512 karakter)
- *instruction* (instruksi yang diberikan untuk *assistant*, maksimal 32.768 karakter)
- *tools* (dapat berisi *code interpreter* (AI akan membuatkan code yang akan di *run* untuk menyelesaikan tugas), *retrieval* atau *function*, maksimal 128 *tools*)
- *file\_ids* (untuk memberikan list id *file* yang di *upload*, berguna jika memakai *tools retrieval*)

*Outputnya* berupa objek *assistant*.

Untuk membuat *thread* baru, dibutuhkan JSON dengan *parameter list object message* yang ada. *Outputnya* akan berupa *object thread*.

Untuk membuat *object message* dibutuhkan JSON dengan *parameter*

- *role* (*user*, *system*, *function*, *assistant*)
- *content* (isi dari *message*)
- *file\_id* apabila ada *file* yang digunakan untuk membuat isi *message* tersebut (biasanya *message* jenis ini, berasal dari respon *assistant*).

#### 2.1.4 Llama 2

Llama 2 merupakan produk *pre-trained* LLM yang dibagi menjadi dua jenis yaitu Llama 2 dan Llama 2 *chat*. Llama 2 sendiri telah di *pre-trained* dan juga telah diberi pengamanan (Touvron. H et al, 2023). Untuk model Llama 2 sendiri di *train* dengan *data - data* publik yang ada dan terdiri dari 70B, 13B dan 7B *parameter*. Untuk *model* Llama 2 *chat*, secara khusus di optimisasi untuk dialog chat sehingga lebih optimal sebagai *chatbot* dan juga terdiri dari 70B,13B dan 7B *parameter* seperti model Llama 2 biasa.

Meta menggunakan data data yang bisa diakses secara publik untuk pre-training datanya. Namun, meta telah melakukan penghapusan data - data yang bersifat pribadi. Untuk *pre-training*, *setting* dan arsitektur model kebanyakan sama dengan *model* pendahulunya yaitu Llama-1. Untuk *fine tuning* yang dilakukan pada *model* Llama 2 adalah sebagai berikut :

- *Supervised Fine-Tuning* (SFT): Dilakukan dengan menggunakan sekitar 27.540 data berkualitas tinggi yang diproses secara manual. Data ini digunakan untuk melatih model agar *model* dapat menjawab *prompt* dengan lebih baik.
- *Reinforcement Learning with Human Feedback* (RLHF): Dilakukan dengan mengumpulkan data kebiasaan manusia, kemudian melatih *reward model* untuk meniru kebiasaan tersebut. *Reward model* ini kemudian digunakan untuk melatih kembali model chat agar sesuai dengan kebiasaan manusia. Teknik - teknik RLHF yang digunakan antara lain PPO dan *Rejection Sampling*.
- *Ghost Attention* (GAttn): Suatu teknik yang dilakukan untuk membuat *model* tetap konsisten terhadap instruksi awal selama *conversation* dengan model berlangsung.

#### 2.1.5 Mistral 7B

Mistral 7B adalah sebuah model LLM dengan 7 miliar *parameter* yang dirancang untuk memiliki kinerja dan efisiensi yang unggul (Jiang, A. Q et al., 2023). Mistral 7B memanfaatkan *grouped query attention* (GQA) untuk inferensi yang lebih cepat, ditambah dengan *sliding window attention* (SWA) untuk secara efektif menangani urutan yang panjangnya berubah-ubah dengan biaya inferensi yang lebih rendah. GQA secara signifikan meningkatkan kecepatan inferensi dan juga mengurangi kebutuhan memori selama *decoding*, sementara SWA dirancang untuk menangani urutan yang lebih panjang dengan lebih efektif pada biaya komputasi yang berkurang. Mistral 7B *ditrain* menggunakan pendekatan *self supervised learning* (SFT) pada data teks yang besar. Mistral 7B diuji pada berbagai *benchmark* seperti

*commonsense reasoning, world knowledge, reading comprehension*, matematika, dan *code generation*. Mistral 7B juga didesain sedemikian rupa supaya mudah di *fine tuning* untuk berbagai pekerjaan.

#### 2.1.6 **Fine Tuning**

*Fine tuning* adalah proses penyesuaian ulang suatu model yang sudah *pre-trained* pada dataset umum untuk meningkatkan kinerjanya pada tugas tertentu. *Fine tuning* memiliki beberapa metode yaitu :

- *Full Fine Tuning* : *fine-tuning* pada keseluruhan model, yang biasa dilakukan pada model berukuran kecil seperti BERT atau RoBERTa.
- *Parameter-efficient Fine-Tuning* (PEFT) : Menambahkan sejumlah kecil *parameter* baru ke dalam LLM yang sudah dilatih sebelumnya dan hanya melatih *parameter* tambahan tersebut untuk meningkatkan kinerja dengan biaya yang lebih rendah.
- *Supervised Fine-Tuning* (SFT): Dilakukan dengan menggunakan data khusus yang sudah dilabeli, umumnya dataset tersebut berbentuk dalam *input output*. Sehingga LLM dapat lebih memahami tugasnya

*Fine tuning* sendiri pada LLM dilakukan agar model dapat melakukan tugas yang spesifik dengan tepat dan juga untuk mengurangi tingkat *AI Hallucination*.(Huang, X, n.d.)

#### 2.1.7 **Low Rank Adaptation (LoRA)**

LoRA adalah metode untuk melakukan adaptasi model skala besar seperti LLM agar lebih efisien dalam penyimpanan dan komputasi. LoRA menggunakan representasi *low-rank* untuk melakukan perubahan pada model tanpa harus memakai semua parameter (Edward Hu et al, 2021). Untuk sistemnya pada awalnya ketika melakukan *fine tuning*, akan diidentifikasi parameter mana saja yang akan digunakan. Untuk perubahan modelnya akan disimpan dalam bentuk representasi *low-rank*. Hal ini akan menyebabkan ketika kita melakukan *fine tuning*, *resource* yang digunakan akan lebih sedikit sehingga tidak boros *resource*.

#### 2.1.8 **Text Data Augmentation**

*Text data augmentation* adalah teknik untuk membuat data dari data yang ada untuk memperluas data *training*. Salah satu tujuan kenapa dilakukan *text data augmentation* adalah untuk mencegah *overfitting* (Shorten C, 2021). Beberapa teknik *text data augmentation* antara lain sebagai berikut :

- *Rule-based augmentation*: Membuat aturan *if else* atau template untuk menghasilkan data baru, seperti penghapusan kata secara acak, penyisipan kata secara acak, penggantian sinonim secara acak.
- *Graph-structured augmentation*: Menambahkan struktur seperti *graph* atau *tree* untuk menemukan transformasi yang mempertahankan label.
- *MixUp augmentation*: Membentuk data baru dengan mencampur data yang sudah ada, misalnya menggabungkan setengah kalimat dari satu data dengan setengah data lainnya.
- *Back-translation augmentation*: Menerjemahkan teks dari satu bahasa ke bahasa lain dan diterjemahkan kembali lagi untuk menghasilkan data yang baru.
- *Style augmentation*: Mengubah gaya penulisan dari satu penulis ke penulis lain untuk menghasilkan variasi baru.
- *Generative augmentation*: Menggunakan model generatif terlatih untuk menghasilkan data baru dari data yang ada.

## 2.2 Tinjauan Studi

Berikut merupakan tinjauan studi yang berkaitan dengan pengaplikasian algoritma evolusi dalam menghasilkan *behavior agent*.

### 2.2.1 *Fine-Tuning Llama 2 Large Language Models for Detecting Online Sexual Predatory Chats and Abusive Texts*

Penelitian ini dilakukan dengan menggunakan model Llama 2 yang di *fine tuning* menggunakan dataset percakapan predator seksual (PAN12) dan dataset teks kasar dalam bahasa Roman Urdu dan Urdu. *Fine tuning* dilakukan dengan metode *Low-Rank Adaptation* (LoRA) untuk meningkatkan efisiensi supaya proses *fine tuning* menjadi lebih cepat. Hasilnya di 3 dataset yang berbeda dan 5 set eksperimen, dengan akurasi sempurna dan F1 dan F0.5 dengan nilai 0.98. Kelebihan dari penelitian ini adalah memanfaatkan kemampuan model Llama 2 yang sudah terlatih dengan data *general* dan dibekali pemahaman kata, lalu disempurnakan lagi dengan fine tuning sehingga hasilnya maksimal untuk pendeteksian predator seksual melalui *chat*.

Untuk kelemahan dari penelitian ini adalah data-data yang digunakan masih didominasi data bahasa Inggris, sehingga jika data bahasa lain yang digunakan untuk menguji

mungkin masih kurang optimal. Perbedaan penelitian ini dengan yang akan diajukan adalah *dataset* yang digunakan adalah dataset bahasa Indonesia dan juga pendekatan *fine tuning*nya untuk tujuan *question and answer* bukan untuk pendeteksian.

### **2.2.2 Implementasi Chatbot “Alitta” Asisten Virtual Dari Balittas Sebagai Pusat Informasi Di Balittas**

Penelitian ini menggunakan Telegram, *Knowledge base* dan API.AI untuk bagian *Natural Language Processing* (NLP). Pada penelitian ini dilakukan pembuatan *Chatbot* yang menggunakan data *knowledge base*, namun agar lebih mudah dipahami oleh user maka data tersebut diproses oleh API.AI. Pengujiannya dilakukan dengan metode *Black Box* yaitu pengujian untuk melihat apakah sistemnya berjalan dengan baik. Kelebihannya adalah *Chatbot* dapat menjawab pertanyaan yang sering ditanyakan tanpa terkendala oleh waktu dan dengan penggunaan API.AI dapat membantu user untuk memahami respon dari *Chatbot* dengan mudah.

Perbedaan penelitian ini dengan penelitian yang akan diajukan adalah pada asisten *Chatbot* nya. Pada penelitian ini *Chatbot* nya hanya berfokus pada *question answering* agar pengguna tidak kesusahan dalam mencari data. Pada penelitian yang akan diajukan , asisten *Chatbot* tidak hanya berfungsi untuk mencari data namun juga menganalisa data yang ada untuk membuat data yang baru karena menggunakan Generative AI.

### **2.2.3 Fine-tuning Large Language Models for Adaptive Machine Translation**

Penelitian ini melakukan *fine tuning* pada *Large Language Model* (LLM) untuk tujuan *translation*. *Model* yang digunakan adalah Mistral 7B, untuk *dataset* yang digunakan adalah *dataset* medis OPUS seperti ELRC, EMEA, SciELO, dan TICO-19 untuk bahasa Spanyol-Inggris. Jumlah *dataset* yang digunakan untuk *training* adalah 10.000 *dataset zero shot prompt* dan 10.000 *dataset one shot prompt*. Untuk data evaluasi diambil 1000 data *random* dari *dataset train* dan diuji dengan 10.000 data.

Teknik *fine tuning* yang digunakan adalah QLoRA, dengan menggunakan *Google Collab Pro* dengan GPU NVIDIA A100. *Model* Mistral 7B yang sudah di *fine tuning* dibandingkan dengan Mistral 7B *base* dan 2 *model* lain yaitu NLBB 3.3B (*model* khusus *translation*) dan ChatGPT 3.5 *Turbo*. Untuk *zero shot*, dengan metrik BLEU, NLBB menghasilkan angka 47.02,

Mistral 7B *base* menghasilkan angka 42.88, ChatGPT menghasilkan angka 44.65 dan Mistral 7B *fine tuned* menghasilkan angka 46.71. Untuk *one shot*, dengan metrik BLEU, NLBB menghasilkan angka 47.42, Mistral 7B *base* menghasilkan angka 47.35, ChatGPT menghasilkan angka 48.34 dan Mistral 7B *fine tuned* menghasilkan angka 49.69. Terjadi peningkatan signifikan setelah di *fine tuning* dan juga dapat menyaingi model yang dikhususkan untuk *translation*.

Perbedaan penelitian ini dengan penelitian yang akan diajukan adalah pada *dataset* dan fungsinya. Pada penelitian ini LLM dikhususkan untuk tugas *translation* dan menggunakan *dataset* bahasa inggris dan spanyol. Pada penelitian yang akan diajukan akan difokuskan pada tanya jawab mengenai kitab Alkitab dengan tujuan membantu studi Alkitab serta menggunakan *dataset* yang berbahasa Indonesia.