

**MODEL *CHATBOT* STANDAR OPERASIONAL PROSEDUR LAYANAN
AKADEMIK MENGGUNAKAN *NATURAL LANGUAGE PROCESSING*
(Studi Kasus: Jurusan Teknik Elektro Universitas Lampung)**

(Skripsi)

Oleh

**KAFANIAL KHAFI
NPM 1755031009**



**JURUSAN TEKNIK ELEKTRO
FAKULTAS TEKNIK
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2024**

**MODEL *CHATBOT* STANDAR OPERASIONAL PROSEDUR LAYANAN
AKADEMIK MENGGUNAKAN *NATURAL LANGUAGE PROCESSING*
(Studi Kasus: Jurusan Teknik Elektro Universitas Lampung)**

Oleh

Kafanial Khafi

**Sebagai Salah Satu Syarat untuk Mencapai Gelar
SARJANA TEKNIK**

Pada

**Program Studi S1 Teknik Elektro
Jurusan Teknik Elektro
Fakultas Teknik Universitas Lampung**



**JURUSAN TEKNIK ELEKTRO
FAKULTAS TEKNIK
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2024**

ABSTRAK

MODEL *CHATBOT* STANDAR OPERASIONAL PROSEDUR LAYANAN AKADEMIK MENGGUNAKAN *NATURAL LANGUAGE PROCESSING* (Studi Kasus: Jurusan Teknik Elektro Universitas Lampung)

By

KAFANIAL KHAFI

Penelitian ini membahas tentang pengembangan *chatbot* yang dapat mempermudah akses layanan informasi di Jurusan Teknik Elektro Universitas Lampung bagi mahasiswa, dengan memanfaatkan *natural language processing* (NLP). *Language model* GPT-2 digunakan untuk melaksanakan berbagai tugas NLP. Permasalahan dalam penelitian ini muncul ketika mahasiswa terlambat mendapatkan informasi mengenai standar operasional prosedur, yang mengakibatkan kesulitan dalam menjalankan rencana yang telah disusun sebelumnya secara optimal. Tujuan penelitian ini adalah untuk membuat model *chatbot* standar operasional prosedur layanan akademik menggunakan NLP di Jurusan Teknik Elektro Universitas Lampung. Metode yang digunakan dalam penelitian ini adalah *artificial intelligence methodology*. Hasil penelitian menunjukkan bahwa GPT-2 yang dilatih menggunakan *dataset* internal masih mengalami masalah halusinasi, sehingga jawaban yang dihasilkan oleh model *chatbot* sering kali tidak akurat. Model *chatbot* yang dikembangkan hanya memiliki akurasi sebesar 20,00% dalam menangani berbagai pertanyaan dengan banyak topik. Oleh karena itu, kinerja model *chatbot* ini belum memadai untuk digunakan sebagai alat bantu standar operasional prosedur layanan akademik dengan NLP.

Kata Kunci: *Chatbot, GPT-2, Natural Language Processing, Large Language Model, AutoTrain*

ABSTRACT

STANDARD OPERATING PROCEDURE CHATBOT MODEL FOR ACADEMIC SERVICES USING NATURAL LANGUAGE PROCESSING (Case Study: The Department of Electrical Engineering, The University of Lampung)

By

Kafanial Khafi

This research discusses the development of a chatbot that can facilitate access to information services in the Department of Electrical Engineering at the University of Lampung for students, using natural language processing (NLP). GPT-2 language model is used to perform various NLP tasks. Problem in this research arises when students experience delays in receiving information about standard operating procedures, leading to difficulties in executing their plans optimally. Objective of this research is to develop a chatbot model for standard operating procedure of academic services using NLP in the Department of Electrical Engineering at the University of Lampung. The methodology used in this research is artificial intelligence methodology. Research results indicate that GPT-2, trained with an internal dataset, still experiences hallucination issues, causing the chatbot's responses to often be inaccurate. The developed chatbot model achieved an accuracy of only 20.00% in handling various questions on multiple topics. Therefore, the performance of this chatbot model is not yet adequate for use as a tool for standard operating procedures in academic services using NLP.

Keywords: *Chatbot, GPT-2, Natural Language Processing, Large Language Model, AutoTrain*

Judul Skripsi

: **MODEL *CHATBOT* STANDAR OPERASIONAL
PROSEDUR LAYANAN AKADEMIK
MENGUNAKAN *NATURAL LANGUAGE
PROCESSING* (Studi Kasus: Jurusan
Teknik Elektro Universitas Lampung)**

Nama Mahasiswa

: **Kafanial Khafi**

Nomor Pokok Mahasiswa : 1755031009

Jurusan

: Teknik Elektro

Fakultas

: Teknik

MENYETUJUI

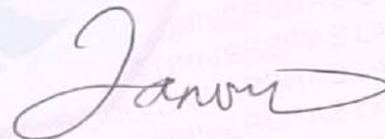
1. Komisi Pembimbing

Pembimbing Utama

Pembimbing Pendamping



Misfa Susanto, S.T., M.Sc., Ph.D.
NIP 19710525 199903 1 001

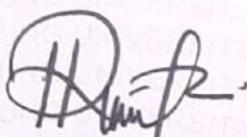


**Ir. Meizano Ardhi Muhammad,
S.T., M.T., IPM.**
NIP 19810528 201212 1 001

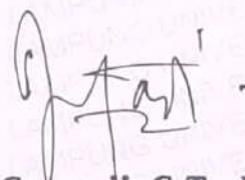
2. Mengetahui

Ketua Jurusan Teknik Elektro

Ketua Program Studi Teknik Elektro



Herlinawati, S.T., M.T.
NIP 19710314 199903 2 001

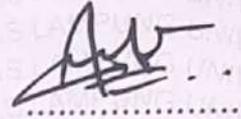


Sumadi, S.T., M.T.
NIP 19731104 200003 1 001

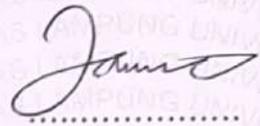
MENGESAHKAN

1. Tim Penguji

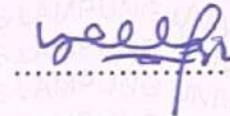
Ketua : **Misfa Susanto, S.T., M.Sc., Ph.D.**



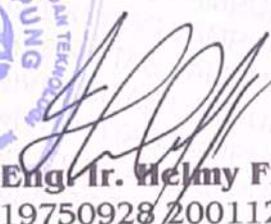
Sekretaris : **Ir. Meizano Ardhi Muhammad,
S.T., M.T., IPM.**



Penguji : **Yetti Yuniati, S.T., M.T.**



2. Dekan Fakultas Teknik

Dr. Eng. Ir. Helmy Fitriawan, S.T., M.Sc. 

NIP 19750928 200112 1 002

Tanggal Lulus Ujian Skripsi : **10 Juni 2024**

SURAT PERNYATAAN

Dengan ini saya menyatakan bahwa dalam skripsi ini tidak terdapat karya yang pernah dilakukan orang lain dan sepanjang sepengetahuan saya tidak terdapat atau diterbitkan oleh orang lain, kecuali secara tertulis diacu dalam naskah ini sebagaimana yang disebutkan dalam daftar pustaka. Selain itu, saya menyatakan pula bahwa skripsi ini dibuat oleh saya sendiri. Apabila pernyataan saya tidak benar, maka saya bersedia dikenai sanksi sesuai dengan hukum yang berlaku.

Bandar Lampung, 12 Juni 2024
Penulis,



Kafanial Khafi

RIWAYAT HIDUP

Penulis dilahirkan pada tanggal 12 Juni 1999 di Tanjung Pinang, dengan Ibu bernama Yuli Ningsih dan Ayah bernama Krisdianto Ari Wibowo. Penulis merupakan anak pertama dari dua bersaudara.

Penulis menyelesaikan pendidikan formal pertama kali di Taman Kanak-Kanak (TK) Barunawati 1 pada tahun 2005, menyelesaikan Sekolah Dasar (SD) di SD Negeri 58 Palembang pada tahun 2011, menyelesaikan Sekolah Menengah Pertama (SMP) di SMP Negeri 8 Palembang pada tahun 2014, kemudian melanjutkan jenjang Sekolah Menengah Atas di SMA YPI Tunas Bangsa Palembang dengan mengambil jurusan Ilmu Pengetahuan Alam (IPA) dan lulus pada tahun 2017.

Pada tahun 2017, penulis terdaftar sebagai mahasiswa Program Studi Teknik Elektro Fakultas Teknik Universitas Lampung melalui jalur SMMPTN. Selama menjadi mahasiswa, penulis pernah bergabung dalam sebuah organisasi yaitu sebagai Anggota Divisi Media Informasi Departemen Komunikasi dan Informasi Himpunan Mahasiswa Teknik Elektro (HIMATRO) pada tahun 2018. Penulis pernah melaksanakan Kerja Praktik (KP) pada bulan April 2021 yang dilaksanakan di PT. Telkom Indonesia, Tbk Regional Palembang pada divisi *Corporate Customer Access Network (CCAN)*.

PERSEMBAHAN

Skripsi ini kupersembahkan untuk

“ Kedua Orangtua”

Yang selalu mendoakan penulis di setiap saat .

Senantiasa memberikan dukungan moril maupun materiil
dalam menyelesaikan Skripsi.

-TERIMA KASIH-

Motto

“Belajarliah menjalani waktu ini, membayangkan waktu di masa depan hanya membuatmu ketakutan, memikirkan waktu pada masa lalu hanya membuatmu kecewa.”

“Jika kamu mengenal musuhmu dan mengenal dirimu sendiri, kamu tak perlu takut pada hasil dari seratus pertempuran. Jika kamu hanya mengenal dirimu sendiri tetapi tidak mengenal musuhmu, untuk setiap kemenangan yang diraih, kamu juga akan menderita kekalahan. Jika kamu tidak mengenal baik musuh maupun dirimu sendiri, kamu akan kalah dalam setiap pertempuran.” – Sun Tzu

SANWACANA

Alhamdulillah rabbil alamin, penulis sampaikan puji syukur ke hadirat Allah SWT yang telah memberikan rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan penelitian Skripsi ini.

Skripsi dengan judul **“MODEL CHATBOT STANDAR OPERASIONAL PROSEDUR LAYANAN AKADEMIK MENGGUNAKAN NATURAL LANGUAGE PROCESSING (Studi Kasus: Jurusan Teknik Elektro Universitas Lampung)”** ini merupakan salah satu syarat untuk memperoleh gelar Sarjana Teknik pada Jurusan Teknik Elektro Fakultas Teknik Universitas Lampung.

Dalam masa perkuliahan dan penelitian, penulis mendapatkan banyak hal baik berupa dukungan, semangat, motivasi, dan banyak hal yang lainnya. Untuk itu penulis mengucapkan terima kasih kepada:

1. Allah SWT atas semua karunia, rahmat serta hidayah-Nya.
2. Nabi Muhammad SAW yang memberikan semangat melalui kisah-kisah yang telah diceritakan.
3. Orangtua dan saudara yang selalu memberikan doa dan semangat dari awal pelaksanaan skripsi hingga terselesaikan.
4. Bapak Misfa Susanto, S.T., M.Sc., Ph.D. selaku pembimbing utama skripsi yang telah dengan sabar dan memberikan begitu banyak waktu untuk membimbing, memberikan ilmu, semangat, motivasi, dan juga arahan.

5. Bapak Ir. Meizano Ardhi Muhammad, S.T., M.T., IPM. selaku pembimbing pendamping skripsi memberikan saran, memberi ilmu motivasi selama masa pendidikan sampai pengujian skripsi.
6. Ibu Yetti Yuniati, S.T., M.T. selaku penguji skripsi yang memberikan kritik dan saran untuk membangun skripsi ini.
7. Bapak Ibu dosen dan pegawai di Jurusan Teknik Elektro yang telah memberikan ilmu, arahan, dan wawasan yang tak terlupakan oleh penulis.
8. Teman-teman HIRO 2017 yang sudah menjadi seperti keluarga sendiri, terima kasih segala kebaikan yang sudah diberikan.

Besar harapan semoga skripsi ini dapat bermanfaat bagi perkembangan ilmu pengetahuan.

Bandar Lampung, 3 Mei 2024



Kafanial Khafi

DAFTAR ISI

DAFTAR ISI	iii
DAFTAR GAMBAR	vi
DAFTAR TABEL	viii
I. PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Tujuan Penelitian	3
1.3 Rumusan Masalah.....	3
1.4 Manfaat Penelitian	3
1.5 Batasan Masalah	4
1.6 Hipotesis	4
1.7 Sistematika Penulisan	4
II. TINJAUAN PUSTAKA	6
2.1 <i>Artificial Intelligence</i>	6
2.2 <i>Knowledge Base</i>	8
2.3 <i>Large Language Model</i>	9
2.4 Universitas Lampung	9
2.4.1 Peraturan Akademik.....	10
2.4.2 Standar operasional prosedur	11

2.5 <i>Natural Language Processing</i>	11
2.6 <i>Chatbot</i>	12
2.7 Hugging Face	13
2.7.1 Autotrain	13
2.7.2 GPT-2.....	14
2.8 Python	16
2.9 <i>Hyperparameters</i>	16
2.9.1 Parameter di <i>Hyperparameters</i>	17
2.9.2 <i>Blackbox Hyperparameter Optimization</i>	18
2.9.2.1 <i>Random Search</i>	18
2.9.2.2 <i>Grid Search</i>	18
2.9.3 Penilaian.....	19
2.9.3.1 <i>Loss</i>	19
2.9.3.2 <i>Gradient Norm</i>	20
2.10 <i>Confusion Matrix</i>	21
2.11 Penelitian Terkait	22
III. METODOLOGI PENELITIAN	32
3.1 Waktu dan Tempat Penelitian.....	32
3.2 Alat dan Bahan.....	32
3.3 Metode Penelitian	33
3.4 Tahapan penelitian	34

3.4.1 <i>Data Preparation</i>	34
3.4.2 <i>Model Training</i>	41
3.4.3 <i>Model Testing</i>	42
3.4.4 Analisis dan penulisan laporan	45
IV. HASIL DAN PEMBAHASAN	47
4.1 Hasil	47
4.1.1. <i>Data Preparation</i>	47
4.1.2 <i>Model Training</i>	54
4.1.3 <i>Model Testing</i>	57
4.1.4 Confusion Matrix	65
4.2 Pembahasan.....	76
V. SIMPULAN DAN SARAN	80
5.1 Kesimpulan	80
5.2 Saran	80
LAMPIRAN	87

DAFTAR GAMBAR

Gambar 2.1 Logo Universitas Lampung	10
Gambar 2.2 Pipeline Natural Language Processing.....	12
Gambar 2.3 GPT-2 Indonesia	15
Gambar 3.1 Flowchart Tahap Penelitian.....	33
Gambar 3.2 Survei Mengenai Pertanyaan yang Sering Ditanyakan di Layanan Akademik JTE Universitas Lampung	35
Gambar 3.3 Salah Satu Contoh Angket yang Telah Diisi dan Angket yang disebarakan Secara <i>Offline</i>	36
Gambar 3.4 Data dari Angket yang Telah Dikumpulkan	37
Gambar 3.5 Surat Edaran yang Membahas Tentang Syarat Administrasi Seminar	40
Gambar 3.6 Format Penulisan <i>Dataset</i>	40
Gambar 3.7 Repositori Hugging Face.....	42
Gambar 3.8 Menjalankan <i>Model Chatbot</i> Menggunakan Google Colaboratory ..	43
Gambar 4.1 Kombinasi 1 dengan <i>learning rate</i> sebesar 0,02 dan <i>epoch</i> sebanyak 500.....	59
Gambar 4.2 Kombinasi 2 dengan <i>learning rate</i> sebesar 0,02 dan <i>epoch</i> sebanyak 1000.....	60
Gambar 4.3 Kombinasi 3 dengan <i>learning rate</i> sebesar 0,002 dan <i>epoch</i> sebanyak 500.....	61

Gambar 4.4 Kombinasi 4 dengan <i>learning rate</i> sebesar 0,002 dan <i>epoch</i> sebanyak 1000.....	61
Gambar 4.5 Kombinasi 5 dengan <i>learning rate</i> sebesar 0,0002 dan <i>epoch</i> sebanyak 500.....	62
Gambar 4.6 Kombinasi 6 dengan <i>learning rate</i> sebesar 0,0002 dan <i>epoch</i> sebanyak 1000.....	63
Gambar 4.7 Kombinasi 7 dengan <i>learning rate</i> sebesar 0,00002 dan <i>epoch</i> sebanyak 500.....	64
Gambar 4.8 Kombinasi 8 dengan <i>learning rate</i> sebesar 0,00002 dan <i>epoch</i> sebanyak 1000.....	64
Gambar 4. 9 Perangkat keras milik Hugging Face	79

DAFTAR TABEL

Tabel 3. 1 Daftar alat yang Digunakan Dalam Penelitian.....	32
Tabel 3. 2 Daftar Bahan yang Digunakan Dalam Penelitian	33
Tabel 3. 3 Daftar Topik Pertanyaan yang Telah Diurutkan	37
Tabel 3. 4 Topik Pertanyaan yang Diambil dari Mayoritas	39
Tabel 3. 5 Kombinasi <i>Learning Rate</i> dan <i>Epoch</i>	41
Tabel 3. 6 Tabel <i>confusion matrix</i>	45
Tabel 4. 1 Topik Peraturan Akademik yang Didapat dari Hasil Survei.....	48
Tabel 4. 2 Topik dan Sumber Peraturan yang Digunakan	49
Tabel 4. 3 Sampel Pertanyaan dan Jawaban	50
Tabel 4. 4 Daftar Kombinasi Parameter.....	55
Tabel 4. 5 Perbedaan Kerugian (<i>loss</i>) saat Pelatihan Terhadap <i>Learning Rate</i> dan <i>Epoch</i>	58
Tabel 4. 7 Parameter yang Digunakan Saat Melatih.....	56
Tabel 4. 8 Sampel Pertanyaan yang Dilakukan Dalam Pengujian <i>Confusion Matrix</i>	66
Tabel 4. 9 <i>Confusion Matrix</i>	71

I. PENDAHULUAN

1.1 Latar Belakang

Pusat layanan informasi adalah layanan yang menyediakan informasi terkait organisasi kepada individu. Jenis-jenis layanan informasi meliputi *helpdesk*, majalah dinding, bagian administrasi, atau kontak yang dapat dihubungi secara tertulis maupun tidak tertulis. Informasi menjadi sangat penting ketika seseorang ingin menjalankan berbagai fungsi dalam organisasi. Salah satu organisasi pendidikan tersebut adalah Jurusan Teknik Elektro [1], yang berada di bawah naungan Universitas Lampung [2]. Jurusan Teknik Elektro memiliki banyak standar operasional prosedur yang mendukung berbagai fungsi di dalamnya.

Permasalahan muncul ketika mahasiswa terlambat mendapatkan informasi tentang standar operasional prosedur. Keterlambatan ini menyebabkan kesulitan dalam menjalankan rencana yang telah disusun sebelumnya secara optimal. Mahasiswa memiliki beberapa opsi untuk mendapatkan informasi. Salah satunya adalah dengan langsung menanyakan kepada bagian administrasi [3], meskipun hal ini bisa memakan banyak waktu karena mahasiswa perlu menempuh jarak yang jauh untuk mencapai bagian administrasi tersebut. Opsi lainnya adalah dengan mengakses situs yang disediakan oleh kampus atau mencari bantuan melalui *helpdesk*. Namun, proses ini juga memakan banyak waktu karena mahasiswa perlu menelusuri berbagai kategori informasi yang tidak umum atau menunggu balasan dari staf.

Teknologi *chatbot* [4] memungkinkan untuk menyimulasikan interaksi percakapan manusia. Sebagai program komputer, *chatbot* digunakan sebagai alat untuk berkomunikasi dengan orang lain. Kemampuan utama *chatbot* terletak pada kemampuannya untuk memberikan respons terhadap pertanyaan. Oleh karena itu, kemampuan *chatbot* ini menjadikannya sebagai pusat layanan informasi. Dengan algoritma yang dapat disesuaikan sesuai dengan kebutuhannya, *chatbot* dapat diprogram untuk memberikan jawaban yang relevan terhadap pertanyaan pengguna.

Sebuah *chatbot* dapat menjadi solusi untuk menyampaikan informasi dari Jurusan Teknik Elektro Universitas Lampung kepada mahasiswa. Ini memungkinkan mahasiswa untuk mengakses informasi yang spesifik tanpa terbatas oleh jarak dan waktu. *Chatbot* berfungsi sebagai pusat layanan yang menggantikan peran manusia dalam menyampaikan informasi. Manusia membutuhkan waktu istirahat, sementara *chatbot* dapat bekerja tanpa henti selama 24 jam. Kehadiran *chatbot* dapat membantu Jurusan Teknik Elektro untuk memberikan informasi tentang standar operasional prosedur secara konsisten. Layanan informasi yang tersedia secara terus-menerus akan memberikan manfaat kepada mahasiswa dalam menjalankan rencana mereka dengan optimal. Akses cepat terhadap informasi menjadi lebih mudah, tanpa perlu memperhatikan kategori-kategori tertentu seperti pada situs web yang mungkin memiliki navigasi yang rumit. Mahasiswa dapat dengan mudah memperoleh informasi yang dibutuhkan dengan bertanya kepada *chatbot*, yang memberikan jawaban terkait secara langsung.

Untuk mempermudah akses layanan informasi di Jurusan Teknik Elektro Universitas Lampung bagi mahasiswa, diperlukan model *chatbot* standar

operasional prosedur layanan akademik menggunakan *natural language processing* [5]. Dengan menggunakan teknologi *natural language processing* (NLP), *chatbot* mampu memahami dan menanggapi pertanyaan dalam bahasa sehari-hari tanpa memerlukan struktur bahasa yang formal.

1.2 Tujuan Penelitian

Tujuan penelitian ini adalah membuat model *chatbot* standar operasional prosedur layanan akademik menggunakan *natural language processing* di Jurusan Teknik Elektro Universitas Lampung.

1.3 Rumusan Masalah

Rumusan masalah pada penelitian ini yaitu:

Bagaimana membuat model *chatbot* standar operasional prosedur layanan akademik menggunakan *natural language processing* di Jurusan Teknik Elektro Universitas Lampung dengan teknologi GPT-2?

1.4 Manfaat Penelitian

Manfaat dari penelitian ini adalah tersedianya model *chatbot* yang mampu menjawab pertanyaan tentang standar operasional prosedur layanan akademik di Jurusan Teknik Elektro Universitas Lampung.

1.5 Batasan Masalah

Batasan masalah pada penelitian ini adalah 10 pertanyaan yang sering ditanyakan, berdasarkan hasil survei tentang standar operasional prosedur di Jurusan Teknik Elektro terkait peraturan akademik Universitas Lampung.

1.6 Hipotesis

Hipotesis dalam penelitian ini adalah model *chatbot* yang mampu menjawab pertanyaan mengenai standar operasional prosedur layanan akademik di Jurusan Teknik Elektro Universitas Lampung.

1.7 Sistematika Penulisan

Adapun sistematika penulisan Penelitian ini adalah sebagai berikut:

I. PENDAHULUAN

Pendahuluan membahas tentang: Latar belakang, tujuan penelitian, rumusan masalah, manfaat penelitian, batasan masalah, dan sistematika penulisan.

II. TINJAUAN PUSTAKA

Tinjauan pustaka membahas tentang: *artificial intelligence; knowledge base; large language model; Universitas Lampung; natural language processing; chatbot; Hugging Face; Python; hyperparameter; penelitian terkait.*

III. METODOLOGI PENELITIAN

Metodologi penelitian memaparkan waktu dan tempat, alat dan bahan, *data preparation, model training, model testing, analisis dan penulisan laporan.*

IV. HASIL DAN PEMBAHASAN

Hasil dan pembahasan memuat tahapan penelitian dan analisis dari hasil pengujian.

V. PENUTUP

Penutup menjelaskan kesimpulan dari penelitian yang dilakukan dan saran yang didasarkan pada hasil data mengenai perbaikan dan perkembangan menuju lebih baik.

II. TINJAUAN PUSTAKA

2.1 *Artificial Intelligence*

Artificial intelligence (AI) [5] adalah tentang penggunaan komputer untuk meniru proses berpikir manusia. Manusia memiliki beragam kemampuan intelektual seperti belajar, memahami, dan membuat keputusan yang rasional. Dalam konteks kecerdasan buatan, komputer diprogram untuk meniru kemampuan tersebut. Kecerdasan buatan dapat dibagi menjadi empat kategori: berpikir seperti manusia, berpikir secara logis, bertindak seperti manusia, dan bertindak secara logis.

Penerapan kecerdasan buatan memberikan banyak keuntungan yang dapat diperoleh di berbagai sektor pekerjaan sesuai dengan tujuan awal pengembangannya [5]. Beberapa manfaat dari penggunaan kecerdasan buatan antara lain:

a. Pekerjaan

Artificial intelligence dapat menggantikan peran manusia dalam menjalankan berbagai jenis pekerjaan. Pekerjaan yang diemban oleh *artificial intelligence* dapat disesuaikan dengan tujuan awal pembuatannya. Semakin banyak data yang digunakan dalam pelatihan, semakin kompleks pekerjaan yang dapat dilakukan oleh *artificial intelligence*.

b. Tidak berhenti

Sebagai program komputer, *artificial intelligence* mampu beroperasi secara terus-menerus selama terhubung dengan pasokan daya dan menerima perawatan yang memadai. Berbeda dengan manusia, *artificial intelligence* tidak memerlukan istirahat atau terganggu oleh kegiatan lainnya.

c. Tidak ada risiko bahaya

Saat ini, *artificial intelligence* tidak menimbulkan risiko bahaya bagi manusia saat mengalami kerusakan. *Artificial intelligence* dapat menjalankan tugas-tugas berisiko tanpa dipengaruhi oleh emosi, sehingga dapat bekerja dengan optimal tanpa risiko yang terkait. Dalam situasi di mana pekerjaan berpotensi berbahaya, *artificial intelligence* dapat menyelesaikannya tanpa kecemasan atau ketakutan.

d. Bertindak sebagai alat bantu

Artificial intelligence dapat digunakan sebagai alat bantu dalam berbagai situasi, termasuk dalam pengawasan keamanan selama 24 jam. Contohnya, penggunaan *artificial intelligence* untuk mengawasi keamanan rumah dari potensi tindak kejahatan atau untuk mengawasi anak yang melakukan aktivitas berbahaya sendirian.

e. Fungsinya tidak terbatas

Artificial intelligence dapat dirancang untuk memenuhi berbagai kebutuhan. Dalam konteks pekerjaan, *artificial intelligence* cenderung membuat sedikit kesalahan dan dapat meningkatkan efisiensi pekerjaan. Hal ini memungkinkan manusia untuk fokus pada tugas-tugas lain sementara *artificial intelligence* mengambil alih tugas-tugas tertentu.

2.2 Knowledge Base

Knowledge base (KB) [6] adalah suatu teknologi komputer yang memainkan peran penting dalam pengembangan *artificial intelligence*. Dalam *artificial intelligence*, *knowledge base* digunakan untuk menyimpan pengetahuan yang dilatih oleh *artificial intelligence*. Data yang disimpan dalam *knowledge base* sangat beragam biasanya berupa pertanyaan dan jawaban yang telah dipersiapkan sebelumnya. Tujuan dari penyimpanan data adalah untuk memudahkan *artificial intelligence* membuat keputusan dalam memberikan jawaban. Tujuan dari penyimpanan data adalah untuk memudahkan *artificial intelligence* membuat keputusan dalam memberikan jawaban, sehingga sistem dapat beroperasi dengan lebih efisien dan efektif.

Knowledge base dapat meningkatkan efisiensi waktu saat *artificial intelligence* melaksanakan tugas. *Knowledge base* akan membantu *artificial intelligence* saat dijadikan sebagai pusat layanan. Dengan adanya *knowledge base*, *artificial intelligence* dapat dengan mudah mengakses informasi yang telah tersedia dan memberikan jawaban secara cepat kepada pengguna. Sebagai contoh, ketika pengguna mengajukan pertanyaan, *artificial intelligence* dapat langsung merujuk ke *knowledge base* dan menyediakan jawaban yang sudah disiapkan sebelumnya. Selain itu, *knowledge base* juga memungkinkan *artificial intelligence* untuk terus belajar dengan menambahkan data pelatihan baru dan memperbarui informasi yang sudah ada. Selain itu, *knowledge base* juga memungkinkan *artificial intelligence* untuk terus belajar dengan menambahkan data pelatihan baru di saat pelatihan selanjutnya dilakukan sehingga menjaga *artificial intelligence* tetap menyesuaikan informasi yang baru.

2.3 Large Language Model

Large language model (LLM) [4] merupakan model bahasa yang dibuat berdasarkan *transformer language model* (TLM) dan memiliki jutaan parameter bahkan miliaran parameter di dalamnya. *Large language model* memiliki kemampuan untuk memahami bahasa manusia karena dibuat berdasarkan *transformer language model* yang memiliki mekanisme untuk memperhatikan kata yang relevan dalam seluruh teks untuk menjawab pertanyaan yang ditanyakan oleh penanya. *Large language model* yang dilatih menggunakan data pelatihan yang besar akan membuat model tersebut mampu memahami berbagai bahasa yang digunakan oleh manusia. Semakin besar *dataset* yang digunakan untuk melatih *large language model*, semakin baik kinerja LLM yang dihasilkan. Jumlah data yang besar memberikan informasi yang diperlukan untuk meningkatkan pemahaman dan respons model terhadap berbagai jenis pertanyaan, sehingga memungkinkan LLM untuk memberikan jawaban yang lebih akurat dan kontekstual.

2.4 Universitas Lampung

Universitas Lampung (Unila) [2] merupakan organisasi pendidikan dan menjadi salah satu dari sekian banyak universitas di Indonesia. Universitas Lampung terletak di Jl. Prof. Dr. Sumantri Brojonegoro di kota Bandar Lampung. Warna hijau merupakan warna yang menjadi ciri khas dari kampus ini karena warna almamater yang digunakan adalah warna hijau. Sebagai institusi pendidikan yang dikenal luas, Universitas Lampung memiliki logo yang merepresentasikan identitasnya. Gambar 2.1 adalah logo yang digunakan oleh Universitas Lampung.



Gambar 2.1 Logo Universitas Lampung

Sebelum Universitas Lampung berdiri sendiri, Unila menjadi cabang dari Universitas Sriwijaya pada 1961. Dua fakultas yang pertama dibangun adalah Fakultas Ekonomi dan Fakultas Hukum di wilayah Lampung. Tanggal 23 September 1965 merupakan awal baru dari Universitas Lampung dan harapan masyarakat Lampung untuk memiliki universitas sendiri di wilayahnya terkabul. Surat dari Menteri Perguruan Tinggi dan Ilmu Pengetahuan (PTIP) Nomor 195 Tahun 1965 menjadi awal mula berdirinya Universitas Lampung. Dua fakultas yang telah berdiri sebelumnya di wilayah Lampung bergabung dengan Universitas Lampung. Sebelum seluruh fakultas dikumpulkan dalam satu wilayah, Universitas Lampung memiliki tiga lokasi yaitu di Jl. Hasanudin, Jl. Jendral Suprpto, dan Jl. Sorong Cimeng. Sekarang seluruh wilayah Universitas Lampung dikumpulkan di Gedongmeneng. Universitas Lampung yang berada di satu wilayah membuat jarak antara fakultas menjadi dekat.

2.4.1 Peraturan Akademik

Peraturan akademik adalah peraturan yang dibuat dengan tujuan untuk dijadikan pedoman oleh sivitas akademik. Universitas Lampung memiliki sebuah pedoman peraturan akademik yang diatur dalam peraturan rektor Universitas Lampung

nomor 12 tahun 2022. Peraturan akademik ini menjadi acuan dalam kegiatan akademik yang berada di Universitas Lampung.

2.4.2 Standar operasional prosedur

Standar operasional prosedur adalah serangkaian tindakan yang terstruktur dan diatur untuk menjalankan berbagai kegiatan. Jurusan Teknik Elektro berada di bawah Universitas Lampung. Sebagai sebuah jurusan, Teknik Elektro menerapkan berbagai standar operasional prosedur yang mengacu pada peraturan akademik yang dimiliki oleh Universitas Lampung.

2.5 *Natural Language Processing*

Natural language processing (NLP) [5] memberikan kemampuan bagi komputer untuk membaca dan memahami bahasa yang digunakan oleh manusia. Banyak peneliti berharap bahwa sistem *natural language processing* memiliki kemampuan untuk mendapatkan pengetahuan secara mandiri di dalam internet. Beberapa penerapan *natural language processing* secara langsung adalah pengambilan informasi dari pertanyaan dan penerjemahan bahasa ke dalam bahasa yang dipahami komputer.

Chatbot diciptakan sebagai program komputer yang menyimulasikan percakapan antara manusia dan komputer. Hal ini membuat *natural language processing* dirancang sebagai teknologi yang mampu untuk memahami dan merespons teks secara mirip dengan interaksi manusia. Sebagai contoh, ketika seseorang berbicara dalam bahasa sehari-hari dan bertanya kepada komputer, kemudian komputer mencoba untuk mengerti apa yang dikatakan orang tersebut, itulah yang disebut

natural language processing. *Natural language processing* adalah cabang ilmu yang berada di ilmu komputer dan ilmu linguistik. Dengan demikian, NLP membahas interaksi antara komputer dan bahasa natural manusia.

Gambar 2.2 menunjukkan proses bagaimana pertanyaan diolah menggunakan *natural language processing*. Tahap pertama adalah *text processing*, yang merupakan kegiatan mengolah dan membersihkan pertanyaan yang diterima. Tahap kedua adalah *feature extraction*, yang berfungsi untuk mengekstraksi dan mengetahui kesamaan kata yang dimiliki di *knowledge base*. Tahap ketiga ialah *modeling*, yang merupakan kegiatan penyesuaian pertanyaan dengan *knowledge base* yang dimiliki. Tahap *modeling* memiliki tujuan untuk memprediksi maksud dari pertanyaan.



Gambar 2.2 *Pipeline Natural Language Processing*

2.6 Chatbot

Chatbot [4] atau biasa disebut sebagai *chatterbot* merupakan sebuah program komputer yang dirancang agar dapat melakukan simulasi percakapan yang dilakukan oleh manusia. Ketika sebuah pertanyaan dikirimkan oleh pengguna, selanjutnya *chatbot* akan menerima pertanyaan yang telah dikirimkan. Kemudian, pertanyaan yang telah diterima oleh *chatbot* diolah sehingga *chatbot* dapat memberikan respons dari pertanyaan sebelumnya. Jika *chatbot* berhasil menjawab

pertanyaan yang diajukan maka *chatbot* dapat dikategorikan sebagai “sistem dialog manusia-komputer dengan bahasa alami”. Keadaan ini merupakan bentuk melakukan simulasi percakapan manusia. Keutamaan dari *chatbot* untuk melakukan tanya dan jawab akan sangat membantu banyak orang ketika mencari informasi. *Chatbot* memiliki berbagai kemampuan sesuai dengan apa yang dibutuhkan dalam bidang tertentu. Contoh penggunaan *chatbot* adalah menggunakan *chatbot* sebagai *customer service* untuk menjawab pertanyaan dari pelanggan. Selain itu, kemampuan *chatbot* sebagai *customer service* adalah mampu untuk melakukan tugas selama 24 jam.

2.7 Hugging Face

Hugging Face [7] ialah platform yang mendukung perkembangan *natural language processing*. Platform ini banyak digunakan oleh peneliti untuk bertukar ide. Peneliti lain dapat berkontribusi dalam proyek terbuka yang dilakukan oleh berbagai pihak. Alat yang dimiliki oleh Hugging Face untuk mengembangkan *natural language processing* adalah AutoTrain. Hugging Face juga menyediakan berbagai model populer yang dapat digunakan secara gratis seperti BERT dan GPT. *Library* yang dikembangkan oleh Hugging Face dengan menggunakan bahasa pemrograman Python dapat membantu dalam pengembangan *Natural Language Processing* menjadi lebih mudah saat melakukan proses pelatihan.

2.7.1 Autotrain

Autotrain [7] merupakan *tool* yang dikembangkan oleh Hugging Face untuk membantu pelatihan *Artificial Intelligence* dan *Natural Language Processing*.

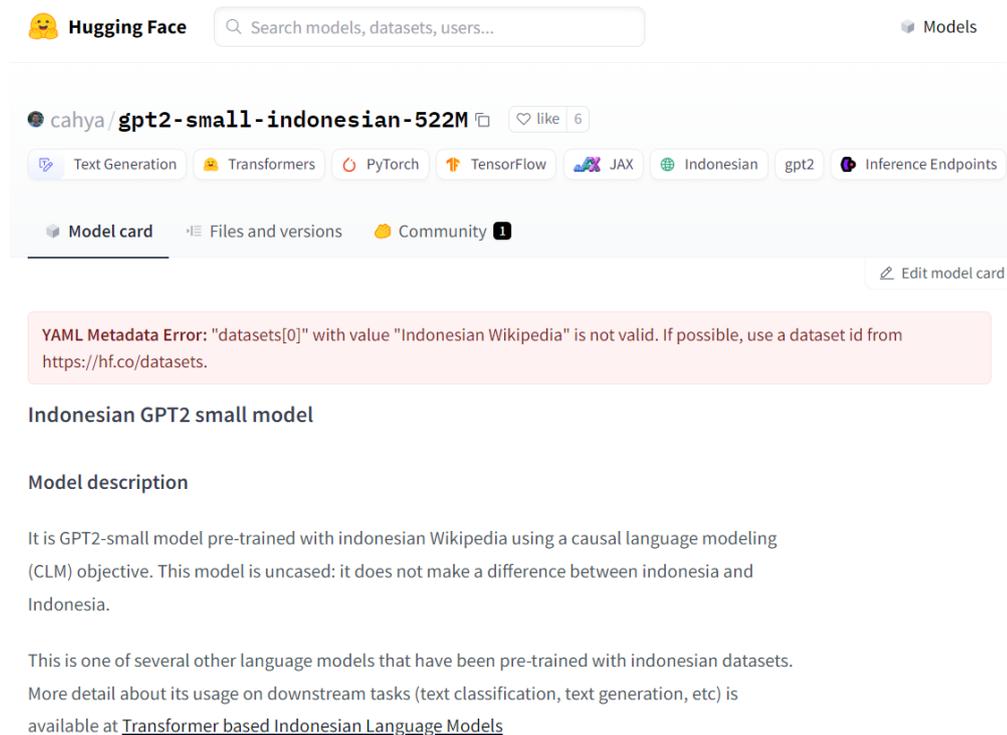
Dalam tahapan melatih AI perlu waktu dan pengaturan data yang rumit. AutoTrain dapat melatih berbagai model yang tersedia pada platform Hugging Face. AutoTrain adalah teknologi yang membantu melatih model dengan efisien serta terintegrasi dalam ekosistem Hugging Face.

Pelatihan yang dapat dilakukan menggunakan AutoTrain adalah LLM *finetuning*, *image classification*, *tabular classification* dan *regression*, *translation question answering*, *summarization*, *entity recognition*, serta *text classification*. Format data yang digunakan juga beragam sesuai dengan keperluan pelatihan model seperti CSV, TSV, JSON, dan ZIP. Selain itu terdapat berbagai cara penulisan sebuah *dataset* yang ditunjukkan dalam instruksi di dokumentasi Hugging Face. Banyaknya pilihan membuat Autotrain dapat digunakan dalam berbagai kondisi sesuai tujuan untuk apa sebuah model akan digunakan.

2.7.2 GPT-2

GPT-2 [8] merupakan salah satu *large language model* yang dikembangkan oleh OpenAI. GPT-2 merupakan penerus dari generasi sebelumnya yaitu GPT. GPT-2 memiliki nama panjang sebagai *Generative Pre-trained Transformer* yang berarti GPT-2 memiliki kemampuan untuk menghasilkan *output*, dilatih dalam beberapa data tertentu untuk memahami struktur kata, dan menggunakan jaringan saraf tiruan berbasis *transformer*. Kelebihan untuk menggunakan GPT-2 adalah GPT-2 dapat digunakan secara gratis serta memiliki ukuran *file* yang kecil. Selain itu, GPT-2 mampu dengan baik untuk menangani tugas dalam *natural language processing* seperti kemampuan untuk menghasilkan teks yang menyerupai manusia. Gambar 2.3 merupakan gambar repositori GPT-2 Indonesia yang berada di platform

Hugging Face. GPT-2 ini telah memperoleh pelatihan dengan *dataset* berbahasa Indonesia.



Gambar 2.3 GPT-2 Indonesia

Versi GPT-2 Indonesia dilatih oleh Cahya Wirawan. GPT-2 Indonesia dilatih dengan *dataset* Wikipedia Indonesia. Wikipedia Bahasa Indonesia yang digunakan memiliki ukuran sebesar 522MB. GPT-2 bahasa Indonesia didapat dari Hugging Face dalam repositori `cahya/gpt2-small-indonesian-522M` [9]. GPT-2 Indonesia memiliki arsitektur dasar GPT-2 versi *small*. GPT-2 versi *small* memiliki 117 juta parameter. Parameter ini merujuk pada bobot dan bias yang dapat disesuaikan selama proses pelatihan. Bobot merupakan nilai-nilai yang menghubungkan interkoneksi antara neuron-neuron yang terletak pada lapisan-lapisan yang berbeda dalam struktur jaringan saraf. Fungsi utamanya adalah menentukan tingkat

signifikansi *input* yang diberikan oleh neuron sebelumnya terhadap pengaktifan neuron berikutnya. Bias merupakan nilai tambahan yang disertakan pada setiap neuron dalam struktur jaringan saraf. Peran utamanya adalah memfasilitasi pembelajaran pola yang lebih rumit oleh jaringan, serta memberikan kesempatan bagi jaringan untuk mengeksplorasi hubungan yang lebih mendalam antara masukan (*input*) dan keluaran (*output*).

2.8 Python

Python [10] merupakan salah satu bahasa pemrograman yang banyak digunakan di dunia. Berbagai alasan Python terkenal adalah mudah untuk dipelajari karena tata bahasa dalam penulisan Python tidak memiliki banyak penggunaan simbol dan dianggap menjadi bahasa yang mudah diingat karena algoritma yang dapat dituliskan sederhana. Python juga dapat digunakan secara gratis dengan mengunduh di situs resmi Python. Banyak *operating system* yang dapat menggunakan Python seperti Windows, Mac, dan Linux. Banyak modul yang telah disediakan oleh Python dan hal ini membantu *programmer* untuk melakukan penerapan Python ke berbagai bidang salah satunya yaitu ilmu data dan *machine learning*.

2.9 Hyperparameters

Hyperparameters [11] adalah cara untuk mengontrol parameter dan dapat disesuaikan untuk pelatihan model sebelum pelatihan model dilakukan. Fungsi pengaturan *hyperparameter* ini dilakukan untuk mengatur proses berjalannya pelatihan model *chatbot*. Menggunakan *hyperparameter* membantu proses

pelatihan berjalan dengan efisien setelah melakukan penyesuaian parameter kepada model yang akan dilatih. Ada beberapa parameter yang dapat dikontrol contohnya *learning rate* dan jumlah *epoch*.

2.9.1 Parameter di *Hyperparameters*

Beberapa parameter yang dapat dikontrol dalam hyperparameter adalah sebagai berikut:

2.9.1.1 *Learning Rate*

Learning Rate [12] adalah salah satu parameter yang dapat dikontrol dalam menjalankan pelatihan model *chatbot*. Dalam pelatihan *chatbot*, *learning rate* digunakan untuk menentukan seberapa besar langkah yang dilakukan. Jika langkah yang diambil terlalu besar maka pelatihan akan menjadi sangat cepat dan *chatbot* yang dihasilkan menjadi tidak stabil. Sebaliknya, jika langkah yang diambil terlalu kecil maka pelatihan akan menjadi sangat lama bahkan bisa menyebabkan pelatihan *chatbot* menjadi tidak bergerak. Dengan demikian, pemilihan *learning rate* yang tepat sangat penting untuk mencapai titik yang optimal dalam pelatihan model.

2.9.1.2 *Epoch*

Epoch [12] adalah salah satu parameter yang dapat dikontrol dalam menjalankan pelatihan model *chatbot*. Dalam pelatihan *chatbot*, *epoch* merupakan keadaan ketika semua data pelatihan telah dilatih sebanyak satu kali putaran. Penggunaan *epoch* dalam parameter adalah menentukan seberapa banyak iterasi yang dilakukan dalam pelatihan *chatbot*. Jika semakin banyak *epoch* yang dilakukan maka *chatbot* akan memahami dengan lebih baik. Oleh karena itu, penting untuk menemukan jumlah *epoch* yang optimal agar model *chatbot* dapat menghasilkan kinerja terbaik.

2.9.2 Blackbox Hyperparameter Optimization

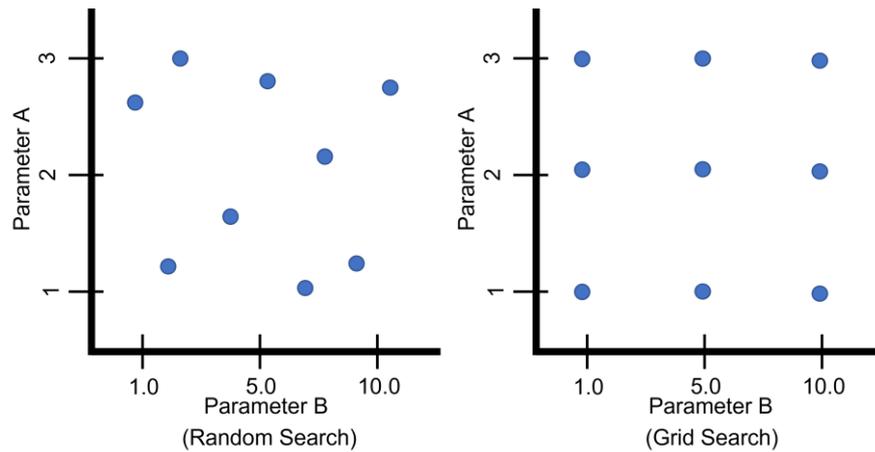
Pendekatan yang dapat digunakan untuk mengoptimalkan proses *hyperparameter* adalah *blackbox hyperparameter optimization* [11]. Metode ini mengoptimalkan *hyperparameter* tanpa asumsi yang dilakukan secara jelas mengenai struktur dari sebuah model. Metode yang digunakan melalui pendekatan *blackbox hyperparameter optimization* adalah *random search* dan *grid search*

2.9.2.1 Random Search

Random search [11] merupakan metode yang digunakan untuk mengoptimalkan *hyperparameter*. Nilai yang digunakan untuk *hyperparameter* dipilih secara acak. Metode ini mirip dengan *grid search* kecuali nilai yang dipilih secara acak. Keunggulan menggunakan *random search* adalah parameter yang digunakan dapat menjadi lebih detail. Tetapi, metode ini menggunakan sumber daya yang banyak karena diperlukan pelatihan yang berulang di setiap titik secara acak.

2.9.2.2 Grid Search

Grid search [11] merupakan metode yang digunakan untuk mengoptimalkan *hyperparameter*. Nilai yang digunakan untuk *hyperparameter* dipilih secara bebas. Keunggulan menggunakan *grid search* adalah dapat mengevaluasi sejumlah besar titik parameter secara efisien. Jadi, penggunaan sumber daya yang digunakan tidak besar. Gambar 2.4 merupakan cara pemilihan parameter yang digunakan untuk mengoptimalkan *hyperparameter*. Metode ini juga memungkinkan untuk menemukan kombinasi *hyperparameter* yang memberikan kinerja terbaik bagi model dengan waktu yang lebih sedikit.



Gambar 2.4 Perbedaan *Random Search* dan *Grid Search*

2.9.3 Penilaian

Penilaian akan dilakukan dengan pendekatan *global minima* [13]. Pendekatan ini adalah menentukan nilai dengan mengamati titik terendah dari keseluruhan dan akan memberikan solusi terbaik dalam pelatihan.

2.9.3.1 Loss

Loss [14] merupakan kerugian yang didapatkan ketika pelatihan *chatbot* dijalankan. Pada penggunaan *grid search*, *loss* digunakan untuk membandingkan kerugian masing-masing kombinasi parameter. Penentuan *loss* dapat dikatakan baik dengan membandingkan nilai kerugian yang dihasilkan *chatbot* saat pelatihan terhadap masing-masing kombinasi parameter. Jika nilai *loss* yang didapatkan saat pelatihan *chatbot* semakin kecil maka *chatbot* akan menjadi lebih baik. *Cross-entropy loss* digunakan dalam menghitung *loss* saat pelatihan. *Loss* dihitung menggunakan persamaan (2.1) [14]:

$$\text{Cross Entropy Loss} = -\frac{1}{N} \sum_{i=1}^N \sum_{j=1}^V y_{ij} \log(p_{ij}) \quad (2.1)$$

Di mana:

N = Jumlah sampel dalam dataset

V = Jumlah *token* atau kosakata yang dimiliki oleh model

y_{ij} = Indikator jika kata berikutnya dari kata ke- i dalam sampel ke- j adalah benar maka *model* akan memberi nilai 1, sebaliknya jika salah maka *model* akan memberikan nilai 0.

p_{ij} = Prediksi model untuk kata ke- i dalam sampel ke- j

2.9.3.2 Gradient Norm

Grad norm atau *gradient norm* [14] pada *hyperparameter* adalah nilai yang mengukur seberapa besar perubahan yang terjadi ketika melatih model. Nilai *gradient norm* dipengaruhi oleh parameter yang diatur ketika melatih. Gradient norm digunakan untuk mengamati stabilitas saat pelatihan. Cara menentukan sebuah kombinasi parameter dapat dikatakan baik adalah dengan menilai perubahan yang terjadi saat latihan dilakukan. *Gradient norm* yang baik akan dengan cepat mencapai kesetimbangan dan tidak terjadi perubahan nilai yang signifikan. Dalam hal ini, jika nilai *gradient norm* mengalami perubahan secara signifikan, maka kombinasi parameter yang digunakan dianggap tidak baik. *Gradient norm* dihitung menggunakan persamaan (2.2) [14]:

$$\text{Gradient norm} = \|\nabla L\| = \sqrt{\left(\frac{dL}{d\theta_1}\right)^2 + \left(\frac{dL}{d\theta_2}\right)^2 + \dots + \left(\frac{dL}{d\theta_n}\right)^2} \quad (2.2)$$

Di mana:

L = Fungsi *loss*

θ = jumlah parameter dalam *language model*, yang terdiri dari bobot (*weights*) dan bias.

2.10 Confusion Matrix

Confusion matrix [15] adalah alat yang digunakan untuk mengevaluasi performa *language model*. *Confusion matrix* membantu untuk memahami tentang sejauh mana *language model* dapat menjawab pertanyaan sesuai dengan nilai yang sebenarnya. Selain itu, penggunaan *confusion matrix* membantu untuk mengenali seberapa besar kesalahan prediksi oleh *language model*. Gambaran visual menggunakan tabel *confusion matrix* memperjelas hasil prediksi *language model* terhadap nilai sebenarnya. Hal ini memungkinkan untuk menilai tingkat akurasi sebuah *language model*. Akurasi adalah metrik yang digunakan untuk mengukur seberapa sering model membuat prediksi yang benar. *Confusion matrix* umumnya berbentuk tabel 2x2 untuk masalah dengan dua kelas yaitu benar dan salah, tetapi penggunaan *confusion matrix* dapat menjadi lebih luas untuk masalah dengan klasifikasi banyak kelas. Gambar 2.5 merupakan contoh penggunaan tabel *confusion matrix* dengan nilai positif dan negatif.

		Nilai Sebenarnya	
		Positif	Negatif
Nilai Prediksi	Positif	TP	FP
	Negatif	FN	TN

Gambar 2.5 Tabel *confusion matrix* untuk nilai positif dan negatif

Persamaan untuk menilai akurasi atau nilai yang diprediksi sesuai dengan hasil dengan menggunakan tabel *confusion matrix* ada pada persamaan (2.3) [15]:

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

Di mana:

True Positive (TP) : Hasil yang diprediksi positif dengan hasil positif

True Negative (TN) : Hasil yang diprediksi negatif dengan hasil negatif

False Positive (FP) : Hasil yang diprediksi positif dengan hasil negatif

False Negative (FN) : Hasil yang diprediksi negatif dengan hasil positif

2.11 Penelitian Terkait

Berikut merupakan penelitian terdahulu yang berhubungan dengan penelitian ini antara lain :

Penelitian [16] yang berjudul “*Dexter the College FAQ Chatbot*” menampilkan arsitektur sistem *chatbot* pada saat ini atau yang lebih dikenal sebagai *retrieval-based chatbot*. Keadaan ini didasari oleh keadaan *chatbot* ketika diberi sebuah pertanyaan. *Chatbot* yang menerima pertanyaan akan mencoba untuk melakukan pencocokan dengan dasar pengetahuannya. Jika *chatbot* memiliki jawaban dari pertanyaan yang ditanyakan maka *chatbot* akan mampu menjawab pertanyaan yang ada. Kala *chatbot* tidak mampu untuk menjawab pertanyaan lalu *chatbot* dapat melakukan panggilan ke admin untuk membantu menyelesaikan pertanyaan. Arsitektur sistem *chatbot* yang diajukan dapat mengatasi masalah admin yang tidak selalu tersedia. Selain itu *chatbot* juga dapat meminta bantuan ketika tidak mampu untuk menyelesaikan masalah yang ditanyakan. Cara kerja dari sistem ini adalah dengan melakukan pemeriksaan ke *database*. *Chatbot* akan memeriksa *database* hingga menghasilkan dua pilihan yaitu jawaban tersedia atau jawaban tidak tersedia. Jika jawaban tersedia maka *chatbot* akan langsung merespons pertanyaan, tetapi jika jawaban tidak tersedia maka *chatbot* akan mencari bantuan. *Chatbot* ini dikembangkan menggunakan *library* dari Rasa. Penggunaan Rasa akan membantu *chatbot* untuk menjawab pertanyaan sesuai konteks. Selain itu, pengembangan *library* Rasa membuat *chatbot* mampu untuk menjaga percakapan tetap pada konteks.

Penelitian [17] yang berjudul “*Natural Language Processing based Jaro-The Interviewing Chatbot*”, mengusulkan proses wawancara dapat dilakukan dengan teknologi *artificial intelligence*. *Chatbot* yang dapat digunakan sebagai pewawancara diberi nama JARO. *Interface* yang digunakan dalam JARO

dikembangkan oleh Google dengan sebuah aplikasi dengan nama Google Dialogflow. JARO akan memiliki tampilan yang mudah untuk digunakan. JARO juga menggunakan aplikasi lain yang dikembangkan dari Google yaitu Google Firebase. Google Firebase adalah aplikasi yang berfungsi sebagai *database* dari JARO. *Chatbot* ini menggunakan *Automatic Question Generation* sehingga *chatbot* mampu untuk membuat pertanyaan yang tidak berulang. Hal ini membantu dalam membuat pertanyaan yang bervariasi dan relevan terhadap pelamar. Kesimpulannya adalah *chatbot* yang didesain dapat membuat proses wawancara menjadi sederhana dan nyaman. Ini juga dapat mempersempit daftar kandidat berdasarkan resume yang dikirimkan yang akan dianalisis berdasarkan kualitas, ketertarikan, dan hobi dari kandidat. Banyaknya pertimbangan yang dilakukan oleh JARO dapat membantu JARO untuk menyiapkan pertanyaan yang relevan dengan kandidat dan dicocokkan dengan kriteria yang dicari oleh perusahaan.

Penelitian [18] yang berjudul “*Development of An e-commerce Sales Chatbot*”, mengusulkan tentang menggabungkan kegiatan penjualan melalui *e-commerce* dan *artificial intelligence*. *Chatbot* yang dibuat akan mampu untuk melakukan kegiatan jual beli selayaknya manusia. Hal ini sangat membantu untuk menembus batas penelitian yang ada pada *natural language processing* dan *natural language understanding*. Cara yang dilakukan adalah dengan membuat *natural language understanding engine* yang difungsikan menjadi server HTTP. Server akan menerima teks sebagai *input* dan membalas teks pesan yang masuk dalam bentuk *JSON string*. Keunggulan lain dari *sales chatbot* adalah *chatbot* mampu merekomendasikan produk menggunakan *recommendation engine*.

Recommendation engine dibuat berdasarkan kriteria pelanggan yaitu produk yang paling diminati oleh pelanggan, jangkauan harga, jenis kelamin, dan pembelian produk sebelumnya. Selain itu *sales chatbot* memiliki keunggulan lain yaitu penggunaan *adaptive pricing engine*. *Adaptive pricing engine* dapat membantu *chatbot* untuk memberikan diskon secara langsung selayaknya manusia. Hasil dari penelitian adalah *natural language understanding engine* dilatih untuk melakukan klasifikasi dari data pelatihan yang disediakan oleh admin. *Chatbot* ini juga dibangun berdasarkan *support vector machine* yang digunakan untuk klasifikasi kata dalam *chatbot*.

Penelitian [19] yang berjudul “*AI and Web-Based Human-Like Interactive University Chatbot (UNIBOT)*”, membuat *University Chatbot (UNIBOT)* yang dapat memberikan respons yang relevan dan efisien. UNIBOT memiliki tampilan yang interaktif dan mampu untuk bekerja dengan memori yang lebih sedikit. Algoritma terbaru yang dikembangkan pada UNIBOT membuatnya menjadi efisien sehingga memerlukan memori yang sedikit untuk operasional. Alur algoritma yang digunakan dalam UNIBOT adalah dengan memanfaatkan pilihan YES atau NO. Algoritma ini muncul ketika pertanyaan yang dikirimkan tidak ditulis dengan baik. Jika YES dipilih maka pengguna akan diberikan beberapa opsi yang mendekati pertanyaannya. Selain itu, jika NO dipilih maka UNIBOT akan menampilkan pernyataan bahwa UNIBOT tidak tahu apa yang ditanyakan. Kesimpulan penelitian UNIBOT adalah dengan memanfaatkan website yang menggunakan *hyper text markup standard*, *cascading style sheet*, dan JQuery dapat menghasilkan tampilan *chatbot* yang interaktif.

Penelitian [20] yang berjudul “*Knowledge Base Collecting Using Natural Language Processing Algorithms*”, berfokus pada penggunaan *natural language processing* untuk memahami data yang bersifat *quasi-structured* atau *unstructured*. Data *quasi-structured* menunjukkan pada data yang memiliki struktur yang tidak sepenuhnya teratur, sementara data *unstructured* adalah data yang tidak memiliki struktur yang jelas. Penggunaan *graph database* dapat membantu untuk melihat hubungan antara bagian-bagian teks dengan jelas. Langkah-langkah yang digunakan adalah mengambil kalimat dari dokumen, memecah kalimat menjadi kata-kata, membangun *graph structure*, dan menyimpan dalam *database*.

Penelitian [21] yang berjudul “*ECHO: A Tool for Empirical Evaluation Cloud Chatbots*” menjelaskan tentang ECHO, sebuah alat evaluasi yang digunakan untuk membandingkan *chatbot* berbasis *cloud* dalam percakapan dengan berbagai bidang yang berbeda. Terdapat tiga *cloud* yang dibandingkan, yaitu IBM Watson, Google Dialogflow, dan Amazon Lex. Dalam melakukan penelitiannya, percakapan dibagi menjadi tiga tingkat, yaitu *basic*, *medium*, dan *complex*. Terdapat dua bidang yang menjadi topik dalam penelitian ini terhadap *chatbot*, yaitu *medical recommendation* dan *flight booking*. Ada tujuh parameter yang digunakan untuk mengevaluasi *chatbot* berbasis *cloud*, yaitu *average response time*, *fallback rate*, *comprehensive rate*, *accuracy*, *precision*, *recall*, dan *F1 score*, yang didapatkan nilainya ketika jawaban *chatbot* benar atau salah. Sebagai hasil dari penelitian evaluasi *chatbot*, *average response time* Amazon Lex lebih unggul dari yang lain di dalam dua bidang yang ditanyakan. Saat *fallback rate*, Amazon Lex juga masih lebih unggul dari dua lainnya. IBM Watson memiliki keunggulan dalam *comprehensive rate* ketika

memperbaiki *input* yang salah oleh pengguna, terutama berkaitan dengan bidang *medical recommendation* dengan percakapan yang kompleks. Dalam hal *accuracy*, Amazon Lex lebih unggul dari yang lain. Saat *precision*, IBM Watson dan Google Dialogflow lebih unggul meninggalkan Amazon Lex. Parameter terakhir adalah F1 *score*, di mana Amazon Lex memiliki skor yang lebih tinggi ketika membicarakan *medical recommendation* dibandingkan dengan yang lain, tetapi ketika membicarakan *flight booking*, kedua lainnya lebih unggul dibandingkan dengan Amazon Lex. Kesimpulan dari penelitian ini adalah ECHO dapat digunakan untuk membandingkan sebuah *chatbot*, yang dapat memudahkan dalam pemilihan platform *chatbot*.

Penelitian [22] yang berjudul “*Chat Bot Using API: Human to Machine Conversation*” menjelaskan keadaan di India tentang penggunaan kereta dan situs kereta yang selalu penuh dikarenakan banyaknya orang melakukan pemesanan. Sebagai alternatif merekomendasikan penggunaan sebuah *chatbot* yang ramah pengguna dan dapat mengakses berbagai informasi terkait kereta. Penggunaan *chatbot* akan menjadi sangat mudah dengan memasukkan sebuah bahasa Inggris yang sederhana dan juga *chatbot* akan terintegrasi dengan berbagai website seperti Facebook, Google Assistant, Viber, dan lainnya. Ada empat tahap dalam mengerjakan penelitian ini yang pertama adalah penggunaan Dialogflow untuk membangun *entities* dan aksi sistem, tahap kedua penggunaan bahasa pemrograman Java untuk membuat *back-end* dari *chatbot*, tahap ketiga menggunakan Postman atau sebuah *website* yang digunakan untuk menguji API, tahap terakhir atau keempat menggunakan situs “<http://railwayapi.com/>” untuk melakukan penyediaan

API agar mendapatkan hasil yang akurat. Salah satu pengujian yang dilakukan adalah dengan membandingkan *chatbot* yang telah dibuat dengan *chatbot* yang memiliki nama MakeMyTrip dan menghasilkan dalam empat tahap pengujian MMT memiliki *response time* yang lebih lama daripada *chatbot* yang telah dibuat dengan nama Railbot. Kesimpulan dari penelitian ini membuat sebuah sistem yang dapat menyediakan informasi tentang kereta dan hal terkait. Dapat digunakan dengan *input* langsung atau berupa teks atau melakukan perintah melalui suara. Dapat menjelajahi berbagai situs kereta yang ada di India hanya dengan satu tempat juga dapat diimplementasikan ke berbagai aplikasi lainnya tanpa perlu melakukan instalasi berbagai aplikasi lainnya.

Penelitian [23] yang berjudul “*Smart College Chatbot Using ML and Python*” menjelaskan pembuatan *chatbot* kampus yang bisa dipasang di *website*. *Chatbot* ini menggunakan *natural language processing* untuk mengerti bahasa manusia dan *machine learning* untuk pelatihan agar bisa memberikan jawaban yang benar. Ketika mengambil data dari *database*, *chatbot* akan mencoba mencocokkan kata kunci. Jika tidak menemukan data, admin diminta untuk menambahkannya. Kesimpulannya, *chatbot* interaktif dan akurat untuk lingkup kampus. Di masa depan, akan diterapkan teknik lain untuk meningkatkan presisinya. Dengan sistem ini, admin bisa melatih *chatbot* dengan pengetahuan lebih luas.

Penelitian [24] yang berjudul “*Towards Higher Customer Conversion Rate: An Interactive Chatbot Using the BEET Model*” menjelaskan *chatbot* yang dibuat

dengan metode *behavioral emotion engagement trigger* (BEET). BEET dapat meningkatkan tingkat konversi pelanggan dengan melakukan pengamatan ke tingkah laku pelanggan. BEET didasari oleh tiga asumsi saat ada empat sifat alami manusia. Asumsi pertama adalah dari empat kepribadian orang hanya memiliki satu sifat yang mendominasi lainnya. Asumsi kedua adalah kemungkinan adanya satu orang memiliki dua sifat yang berdekatan satu sama lain. Asumsi ketiga adalah orang yang berada di luar dari model sifat tersebut. Proses yang dilakukan untuk mendapatkan penilaian sifat pelanggan adalah dengan memisahkan tiap kata melalui proses *natural language processing*. Hasil pemisahan ini menghasilkan kata-kata yang dinilai sebagai acuan dalam menentukan sifat pelanggan. Selain itu, performa *chatbot* dapat dinilai melalui akurasi dalam model BEET dan analisis sentimen dari pelanggan. Tujuan dari dibuatnya BEET adalah meningkatkan kecerdasan emosional dalam percakapan *artificial intelligence* berdasarkan dua aspek, yaitu analisis sentimen dan evaluasi model BEET.

Penelitian [25] yang berjudul “*Infini – A Keyword Recognition Chatbot*” menjelaskan bagaimana cara sebuah *keyword recognition chatbot* bekerja dari pertanyaan yang diajukan dan memberikan jawaban yang sesuai dari basis data. Penelitian *keyword recognition chatbot* menggunakan pendekatan pemrosesan *natural language processing* untuk mengenali kata kunci dalam pertanyaan dan memberikan respons yang sesuai. Proses kerja *keyword recognition chatbot* mencakup dua situasi. Situasi pertama adalah ketika *chatbot* dapat mengidentifikasi pertanyaan, maka akan memberikan jawaban langsung. Situasi kedua adalah jika

chatbot tidak dapat menjawab pertanyaan, respons yang menyatakan tidak memiliki jawaban akan dikirimkan ke pengguna.

Penelitian [26] yang berjudul “*Children Privacy Identification System in LINE Chatbot for Smart Toys*” menjelaskan pembuatan sebuah *chatbot* yang dapat melindungi privasi anak-anak karena *smart toys*. *Smart toys* dapat menyimpan informasi anak. *Chatbot* dapat membantu untuk menghindari dari penyalahgunaan informasi yang semakin marak. Pengembangan CPI dapat diringkas menjadi lima kegiatan yaitu mengunduh kebijakan yang dikeluarkan oleh *smart toys*, melakukan filter terhadap kata yang sensitif, membuat data percakapan dalam format JSON/YAML, membuat *chatbot* dan memasukkan percakapan ke dalamnya dengan menggunakan bahasa program Python, mengunggah data ke *server* Heroku, dan yang terakhir menghubungkan API. Dalam merancang *chatbot* yang aman, perlu adanya algoritma yang memastikan keamanan anak-anak. Algoritma yang digunakan dalam CPI adalah dengan memastikan keamanan data yang dimasukkan ke *smart toys*. Setelah data dimasukkan, *smart toys* akan mengirimkan data tersebut ke *children’s online privacy protection rule* (COPPA) dan *personally identifiable information* (PII). Dua organisasi tersebut dapat membantu dalam pengamanan *smart toys* yang digunakan anak-anak. Algoritma yang dibuat diajukan untuk melindungi anak-anak dari penyalahgunaan informasi.

Penelitian [27] yang berjudul “*The Application of Artificial Intelligence in Foreign Language Teaching*” menjelaskan tentang pengaplikasian *artificial intelligence*

dalam ajar mengajar bahasa asing. Ada beberapa poin pengaplikasian yaitu pengintegrasian *artificial intelligence* dan pengajaran bahasa asing, model untuk pembelajar dan pengajar di bawah lingkup *artificial intelligence* untuk pelajaran bahasa asing, dan perkiraan perkembangan *artificial intelligence* di masa mendatang. Pemodelan yang dilakukan adalah *artificial intelligence* berfungsi menyimpan data seperti gambar, bahasa, ekspresi, kata. Dengan memanfaatkan data yang diambil dari pembelajar maka pengajar dapat menerapkan strategi yang sesuai dengan yang dibutuhkan oleh pembelajar.

Penelitian [28] yang berjudul “*Information Processing and Retrieval from CSV File by Natural Language*” menjelaskan tentang model yang dapat digunakan untuk mengambil informasi di *file* yang memiliki format CSV dengan *natural language*. Langkah dari penelitian ini adalah pertama melakukan *data collection* dengan menyiapkan tiga jenis data yaitu pesan dengan bahasa sehari-hari dalam bentuk formal dan tidak formal dari banyak sumber, *natural language statements* yang berfungsi meningkatkan koleksi bahasa dan data tersebut akan dilatih ke dalam model, dan CSV yang akan digunakan untuk penelitian dan evaluasi. Kedua adalah membuat model yang akan digunakan untuk penelitian dengan pembuatan kamus pada model, pengembangan program, dan melakukan pembaruan model.

III. METODOLOGI PENELITIAN

3.1 Waktu dan Tempat Penelitian

Penelitian dikerjakan pada waktu dan tempat sebagai berikut:

Waktu : Agustus 2023 – Januari 2024

Tempat : Laboratorium Teknik Telekomunikasi, Jurusan Teknik

Elektro, Fakultas Teknik, Universitas Lampung.

3.2 Alat dan Bahan

Alat dan bahan yang digunakan dalam penelitian ini adalah sebagai berikut:

a. Alat

Tabel 3.1 Daftar alat yang Digunakan Dalam Penelitian

No.	Alat	Spesifikasi	Keterangan
1	Laptop Acer V3-471G	Processor Intel® Core™ i5-3230M CPU @ 2.6 GHz, SSD 120GB, RAM 12GB.	Digunakan untuk menjalankan <i>chatbot</i> dan menulis hasil penelitian.
2	Sistem Operasi	Windows 10 Home 64-bit	Digunakan untuk menjalankan laptop.
3	Python	Python 3.11.3	Digunakan sebagai bahasa yang membangun <i>chatbot</i> .
4	Google Colaboratory	Nvidia Tesla T4	Digunakan untuk melatih <i>chatbot</i> .

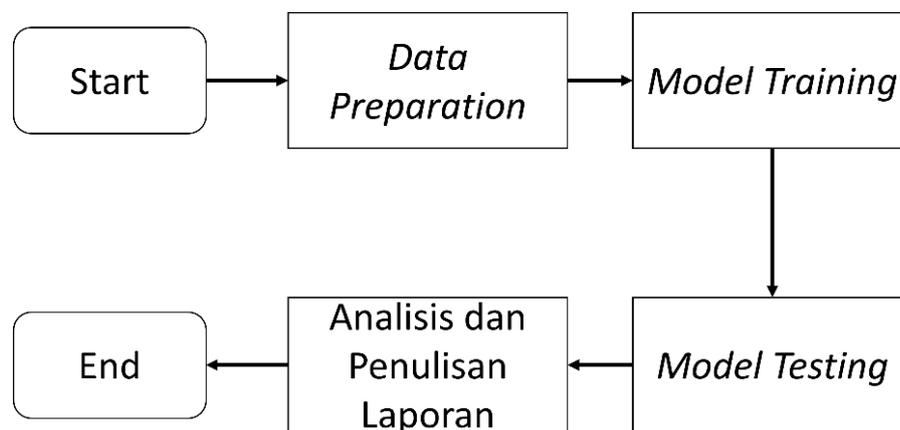
b. Bahan

Tabel 3.2 Daftar Bahan yang Digunakan Dalam Penelitian

No.	Bahan	Keterangan
1	Peraturan akademik Universitas Lampung	Peraturan akademik Universitas Lampung merupakan peraturan yang diterapkan di Universitas Lampung dan menjadi acuan dasar untuk menerapkan standar operasional prosedur di Jurusan Teknik Elektro.
2	Standar operasional prosedur Jurusan Teknik Elektro	Standar operasional prosedur Teknik Elektro merupakan data yang panduan standar operasional prosedur yang ada di Jurusan Teknik Elektro.
3	<i>Language Model GPT-2</i>	GPT-2 merupakan <i>language model</i> yang akan digunakan sebagai objek pelatihan <i>dataset</i> .

3.3 Metode Penelitian

Metode penelitian yang digunakan pada penelitian ini adalah *Artificial Intelligence Methodology*. Detail dari tahapan penelitian ini dapat dilihat dari diagram alir pada Gambar 3.1.



Gambar 3.1 Flowchart Tahap Penelitian

Gambar 3.1 menunjukkan tahapan yang dilakukan dalam penelitian dan penelitian dilakukan dalam empat tahap. Tahap pertama dalam penelitian dimulai dengan

melakukan *data preparation*. Tahap kedua yang dilakukan adalah melakukan *model training*. Kemudian hal yang dilakukan adalah *model testing*. Tahapan terakhir dari tahap penelitian adalah analisis dan penulisan laporan.

3.4 Tahapan penelitian

Detail tentang tahapan penelitian yang dilakukan yang dituliskan pada Gambar 3.1 dijelaskan sebagai berikut.

3.4.1 Data Preparation

Tahapan *preparation* merupakan tahapan awal yang dilakukan dalam penelitian ini. Di dalam tahap *preparation* pertama dimulai dari melakukan survei untuk menemukan 10 standar operasional prosedur yang sering ditanyakan oleh mahasiswa di Jurusan Teknik Elektro.

Pada tahap ini dimulai dengan menggunakan Google Form untuk membuat angket yang dapat disebarakan secara daring dan luring. Penyebaran angket dengan memanfaatkan Google Form dapat menjadi efektif karena angket yang dibuat melalui Google Form dapat disebarakan secara daring dan juga dapat dicetak sehingga dapat disebarakan secara luring. Gambar 3.2 merupakan tampilan aplikasi Google Form yang akan digunakan untuk membuat angket secara daring. Angket dibuat dengan pertanyaan mendasar tentang hal apa saja yang sering ditanyakan pada administrasi Jurusan Teknik Elektro Universitas Lampung. Google Form yang telah dibuat dan disebarakan secara daring kemudian dapat dicetak untuk membantu penyebaran angket secara luring.

Questions Responses **22** Settings

Pertanyaan yang Sering Ditanyakan di Layanan Akademik Jurusan Teknik Elektro Universitas Lampung

B *I* U ↻ ✕

Halo, saya Kafanial Khafi dari jurusan Teknik Elektro UNILA angkatan 2017. Demi melanjutkan penelitian saya, pertanyaan apa saja yang sering kalian tanyakan seputar akademik di jurusan Teknik Elektro?

Nama

Short answer text

NPM

Short answer text

Gambar 3.2 Survei Mengenai Pertanyaan yang Sering Ditanyakan di Layanan Akademik JTE Universitas Lampung

Angket survei yang telah dibuat secara daring melalui Google Form dapat digunakan kembali untuk keperluan menyebarkan survei secara luring. Angket survei tersebut dapat dicetak dalam bentuk fisik. Gambar 3.3 merupakan foto angket yang disebar secara luring di sekitar gedung Jurusan Teknik Elektro Universitas Lampung.

Gambar 3.4 merupakan foto data di dalam aplikasi Microsoft Excel yang sebelumnya telah dikumpulkan dan dibagi ke dalam beberapa kategori. Selanjutnya data-data yang telah dikumpulkan akan dijumlahkan dan diurutkan dari pertanyaan yang sering ditanyakan hingga ke yang jarang ditanyakan. Tabel 3.3 merupakan data dari survei yang telah dikumpulkan dan dibagi menjadi 33 topik pertanyaan

yang sering ditanyakan ke administrasi Jurusan Teknik Elektro Universitas Lampung dan setelahnya survei tersebut dijumlahkan dan diurutkan dari nomor 1 yang merupakan pertanyaan yang paling sering ditanyakan hingga ke nomor 33 yang merupakan pertanyaan yang jarang ditanyakan.

Pertanyaan yang Sering Ditanyakan di Layanan Akademik Jurusan Teknik Elektro Universitas Lampung

Halo, saya Kafanial Khafi dari Jurusan Teknik Elektro UNILA angkatan 2017. Demi melanjutkan penelitian saya, pertanyaan apa saja yang sering kalian tanyakan seputar akademik di jurusan Teknik Elektro?

1. Nama
Annisa Tri Hapsari

2. NPM
2215061025

3. Angkatan
Check all that apply.
 2022
 2021
 2020
 2019
 2018
 2017

4. Pertanyaan apa yang sering kalian tanyakan dengan admin atau Layanan Akademik di jurusan Teknik Elektro? (Contoh: cara mengurus KP, pendaftaran seminar, dll)
Cara mendapatkan cap untuk memproses transkrip nilai

This content is neither created nor endorsed by Google.

Google Forms

Gambar 3.3 Salah Satu Contoh Angket yang Telah Diisi dan Angket yang disebarakan Secara *Offline*

	A	B	C	D	S B U
		Pertanyaan	Syarat Pendaftaran Wisuda	Alur Pendaftaran Wisuda	
2		Nama			
3		Rahmat Fajar Yanto	✓	✗	
4		Dendi Kurniawan	✓	✓	
5		Muhammad Abdul Fattah	✓	✗	
6		Dithiar Prastya	✗	✗	
7		Dzihan Septiangraini	✓	✗	
8		Akmal Dwiki Wijaya	✗	✗	
9		Ni Putu Tiara Prezilia A	✗	✗	
10		Rafif Musyaffa	✗	✗	
11		Irgi Maulana	✗	✗	
12		Fikri Achamadi	✓	✓	
13		M. Fadila Umar	✗	✗	
14		Muhammad Rafi Ramadhan	✗	✗	
15		Irsandi Age Pangestu	✗	✗	
16		Dira Sanjaya Wardana	✓	✗	
17		Ahmad Luthfi Gustomi	✓	✓	
18		Manzalal Aqrom	✗	✗	
19		Ahmad Aldi	✗	✗	
20		Kadafi Eka Sakti	✓	✓	
21		Dzihan Septiangraini	✓	✓	
22		Alizar Abdillah	✗	✗	
23		Ricky Supriadi Sipahutar	✗	✗	
24		Malik Pangestu	✓	✓	
25		Annisa Tri Hapsari	✗	✗	
26		Adinda Ayu Puspitaningrum	✗	✗	
27		Azra Ramadhan Pohan	✗	✗	
28		Gilbran Kadafi	✗	✗	
29		Joko Sulistiyo	✗	✗	
30		Septa Maulana	✗	✗	
31		Faisal Amri	✗	✗	
32		Muhammad Bani Adam	✗	✗	
33		Muhammad Rifqi Majid	✗	✗	
34		Vebry Setia Budi	✗	✗	
35		A. Gilang Alevusta Savada	✗	✗	
36		Aria Seno	✗	✗	
37		Ezra Taufiqurrahman	✓	✗	
38		Bagus Agung Nugroho	✗	✗	
39		Rischo Alfredo D.P	✗	✗	
40		TOTAL	11	6	
41					

Gambar 3.4 Data dari Angket yang Telah Dikumpulkan

Tabel 3.3 Daftar Topik Pertanyaan yang Telah Diurutkan

No.	Pertanyaan	TOTAL
1	Syarat Seminar Proposal	14
2	Syarat Seminar Hasil	13
3	Syarat Komprehensif	13
4	Syarat Seminar KP	12
5	Alur Daftar Seminar	12
6	Alur Pengisian KRS	12
7	Syarat Pendaftaran Wisuda	11
8	Alur Pendaftaran Wisuda	6

Tabel 3.3 Daftar Topik Pertanyaan yang Telah Diurutkan (Lanjutan)

9	Alur Pendaftaran KP	4
10	Alur TTD Transkrip	4
11	Syarat Daftar Beasiswa	4
12	Syarat Banding UKT	3
13	Alur Banding UKT	3
14	Alur Keringanan UKT	3
15	Syarat Pengisian KRS	3
16	Syarat Surat Aktif Kuliah	3
17	Alur Buat Surat Aktif Kuliah	3
18	Syarat Ikut Kampus Merdeka	2
19	Alur Mengurus <i>Learning Agreement</i>	2
20	Syarat <i>Upload Jurnal</i>	2
21	Alur <i>Upload Jurnal</i>	2
22	Syarat Cuti Kuliah	2
23	Alur Mengurus Cuti Kuliah	2
24	Alur Memilih Mata Kuliah	2
25	Alur Menghapus Mata Kuliah	2
26	Alur Peminjaman Alat Lab	2
27	Alur Peminjaman Ruang Seminar	2
28	Alur Pengunduran Diri	2
29	Syarat Kelulusan	2
30	Syarat Keringanan UKT	1
31	Alur Pendaftaran Kampus Merdeka	1
32	Syarat Surat Masih Kuliah	1
33	Alur Buat Surat Masih Kuliah	1

Topik peraturan akademik yang berlaku di Jurusan Teknik Elektro Universitas Lampung dipilih berdasarkan sepuluh hal yang paling sering ditanyakan oleh mahasiswa kepada administrasi. Pertanyaan-pertanyaan tersebut mencakup Syarat Seminar Proposal, Syarat Seminar Hasil, Syarat Komprehensif, Syarat Seminar Kerja Praktek (KP), Alur Pendaftaran Seminar, Alur Pengisian KRS, Syarat Pendaftaran Wisuda, Alur Pendaftaran Wisuda, Alur Pendaftaran KP, dan Alur Tanda Tangan (TTD) Transkrip. Topik-topik ini umumnya dibutuhkan oleh

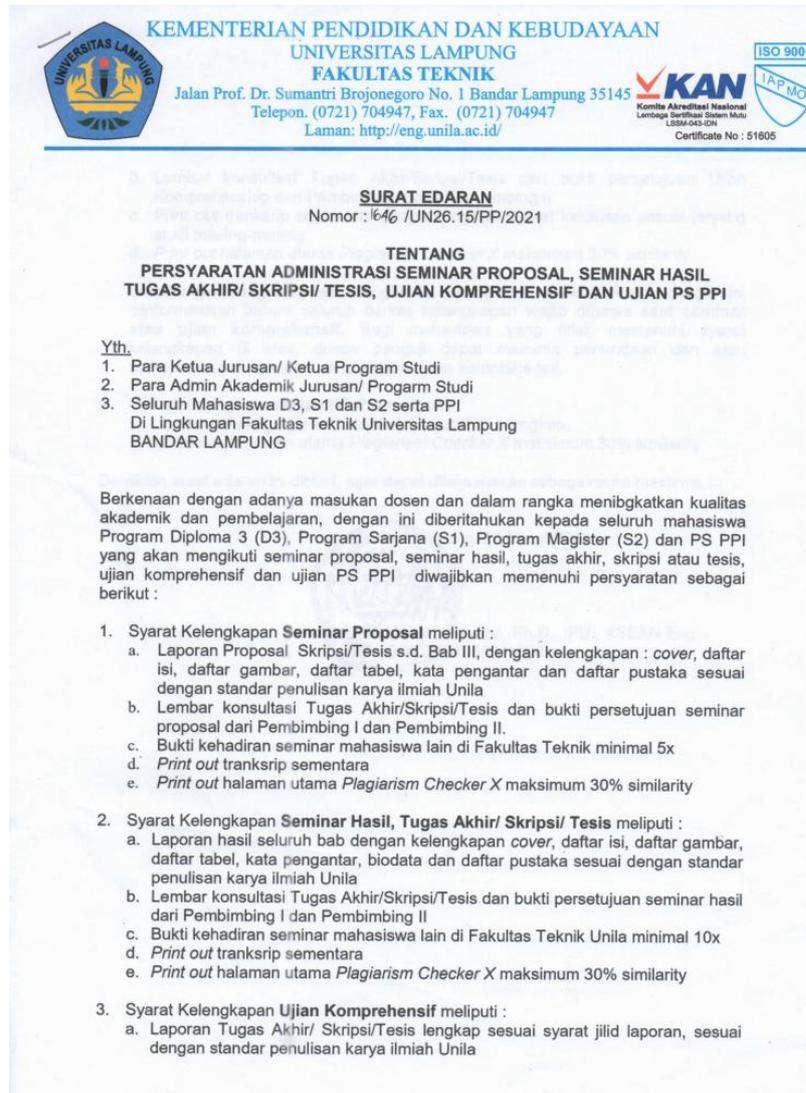
mahasiswa tingkat akhir, mulai dari semester lima hingga akhir. Sepuluh topik utama tersebut dapat dilihat pada Tabel 3.4.

Tabel 3.4 Topik Pertanyaan yang Diambil dari Mayoritas

No.	Pertanyaan
1	Syarat Seminar Proposal
2	Syarat Seminar Hasil
3	Syarat Komprehensif
4	Syarat Seminar KP
5	Alur Daftar Seminar
6	Alur Pengisian KRS
7	Syarat Pendaftaran Wisuda
8	Alur Pendaftaran Wisuda
9	Alur Pendaftaran KP
10	Alur TTD Transkrip

Setelah menentukan sepuluh pertanyaan yang paling sering diajukan oleh mahasiswa di Jurusan Teknik Elektro Universitas Lampung, langkah selanjutnya adalah menemukan peraturan-peraturan yang mengatur topik-topik tersebut. Gambar 3.5 menunjukkan salah satu peraturan dari Fakultas Teknik yang mengatur syarat pendaftaran seminar di tingkat jurusan.

Format penulisan *dataset* dipisahkan oleh tanda ###. Kemudian, ada dua bagian data yang pertama adalah (Human:) yang mewakili pertanyaan dan kemudian komputer (Assistant:) yang mewakili jawaban dari pertanyaan. Semua pasangan pertanyaan dan jawaban berada dalam satu baris yang sama. Satu baris mewakili satu data. *Dataset* yang telah disusun akan digunakan pada tahapan selanjutnya. Detail tentang format yang digunakan dalam penyusunan *dataset* dapat dilihat di Gambar 3.6.



Gambar 3.5 Surat Edaran yang Membahas Tentang Syarat Administrasi Seminar

```

### Human: [Pertanyaan 1] ### Assistant: [Jawaban 1]
### Human: [Pertanyaan 2] ### Assistant: [Jawaban 2]
### Human: [Pertanyaan 3] ### Assistant: [Jawaban 3]
.
.
.
### Human: [Pertanyaan n] ### Assistant: [Jawaban n]

```

Gambar 3.6 Format Penulisan *Dataset*

3.4.2 Model Training

Di dalam tahap *training* akan dilakukan pelatihan kepada GPT-2 bahasa Indonesia. GPT-2 dilatih menggunakan *dataset* yang telah disusun saat tahap *data preparation*. Tahap *model training* memungkinkan model *chatbot* untuk mengidentifikasi pertanyaan. Tahap *model training* menggunakan AutoTrain. Hal ini merupakan alat yang dikembangkan oleh Hugging Face untuk melatih *large language model* dengan menggunakan *dataset* sendiri. AutoTrain menggunakan *hyperparameter* dalam proses melatih *large language model*. *Hyperparameter* diperlukan dalam pelatihan *large language model* agar pelatihan dijalankan secara optimal.

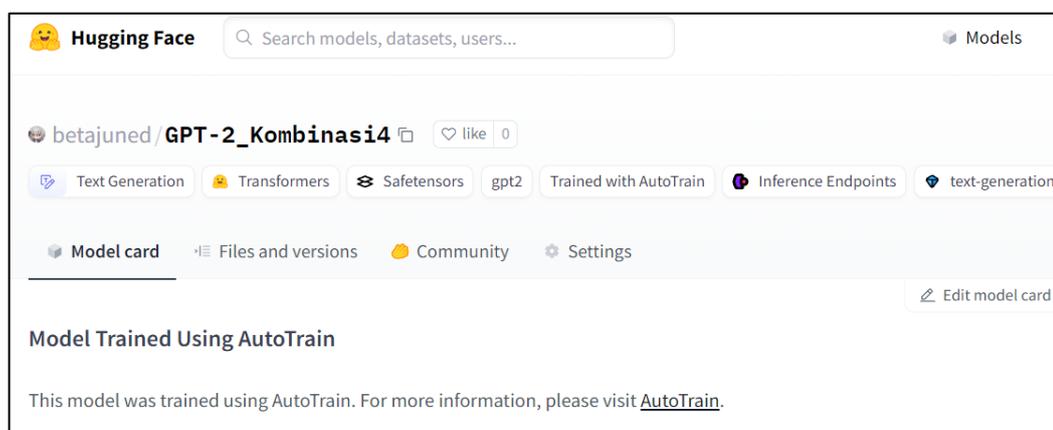
Hyperparameter adalah pengaturan parameter yang digunakan untuk mengontrol proses pelatihan sehingga pelatihan yang dilakukan dapat berjalan dengan optimal. Penentuan parameter dilakukan dengan pendekatan *blackbox hyperparameter optimization* (BHO) yang melihat *language model* sebagai kotak yang tidak diketahui. Salah satu metode adalah *grid search*. Metode *grid search* menggunakan pencarian yang teratur dari beberapa parameter. Penelitian ini mencari parameter dengan mengatur nilai *learning rate* dan *epochi*. *Learning rate* yang digunakan sebanyak 4 variasi. *Epoch* yang digunakan sebanyak 2 variasi. Jadi, terdapat 8 kombinasi dari gabungan antara *learning rate* dan *epoch*. Tabel 3.5 merupakan detail kombinasi parameter yang akan digunakan dalam tahap *model training*.

Melatih *model chatbot* dan mencari parameter yang optimal untuk *model chatbot* dilakukan secara bersamaan. Kemudian, setelah *training model* dilakukan maka *model chatbot* yang telah dilatih disimpan di dalam repositori Hugging Face.

Gambar 3.7 merupakan tampilan repositori yang ada di Hugging Face. Tahap *model training* menghasilkan dua nilai yaitu *loss* dan *gradient norm*. Dua nilai tersebut akan dievaluasi pada tahap *model testing*.

Tabel 3.5 Kombinasi *Learning Rate* dan *Epoch*

Nama	<i>Learning Rate</i>	<i>Epoch</i>
Kombinasi 1	0,02	500
Kombinasi 2	0,02	1000
Kombinasi 3	0,002	500
Kombinasi 4	0,002	1000
Kombinasi 5	0,0002	500
Kombinasi 6	0,0002	1000
Kombinasi 7	0,00002	500
Kombinasi 8	0,00002	1000



Gambar 3.7 Repositori Hugging Face

3.4.3 Model Testing

Tahapan *testing* merupakan tahapan ketiga yang dilakukan dalam penelitian ini. Di dalam tahap *testing* akan dilakukan evaluasi terhadap model yang telah dilatih di tahap *model training*. *Model testing* melakukan evaluasi bertujuan untuk menilai performa GPT-2 untuk digunakan sebagai *model chatbot* yang memiliki data

tentang standar operasional prosedur. Selain itu, melakukan evaluasi juga bertujuan untuk melihat adanya potensi masalah pada *model chatbot*.

Model chatbot dijalankan melalui Google Colaboratory. Kode program yang digunakan untuk menjalankan *model chatbot* berasal dari AutoTrain. Gambar 3.8 merupakan keadaan *model chatbot* saat dijalankan. Kotak berwarna oranye adalah lokasi untuk memasukkan pertanyaan, sedangkan kotak berwarna hijau adalah hasil jawaban yang dikeluarkan oleh *model chatbot*.

```

from transformers import AutoModelForCausalLM, AutoTokenizer

model_path = "betajuned/GPT-2_Kombinasi4"

tokenizer = AutoTokenizer.from_pretrained(model_path, padding_side='left')
model = AutoModelForCausalLM.from_pretrained(
    model_path,
    device_map="auto",
    torch_dtype='auto'
).eval()

# Prompt content: "hi"
messages = [
    {"role": "human", "content": "syarat yang diperlukan untuk seminar hasil?"}
]

input_ids = tokenizer.apply_chat_template(conversation=messages, tokenize=True, add_generation_prompt=True)
output_ids = model.generate(input_ids.to('cuda'), pad_token_id=tokenizer.eos_token_id)
response = tokenizer.decode(output_ids[0][input_ids.shape[1]:], skip_special_tokens=True, max_length=100)

# Model response: "Hello! How can I assist you today?"
print(response)

```

Pertanyaan

Jawaban

A decoder-only architecture is being used, but right-padding was detected! For correct generation, please set padding_side='right' when initializing the tokenizer. (1) Log book pembimbing pertama dan pembimbing

Gambar 3.8 Menjalankan *Model Chatbot* Menggunakan Google Colaboratory

Evaluasi pertama yang dilakukan adalah melihat adanya masalah di kombinasi parameter yang digunakan. Evaluasi ini menilai hasil *loss* dan *gradient norm* setelah pelatihan telah selesai. *Loss* adalah kerugian yang didapatkan saat pelatihan. Tujuan melakukan evaluasi nilai *loss* adalah mengoptimalkan parameter agar nilai *loss* menjadi lebih kecil daripada kombinasi yang lain. Kemudian, menilai *gradient*

norm dari kombinasi parameter yang sudah diatur. *Gradient norm* adalah nilai yang mengukur seberapa besar perubahan yang terjadi saat pelatihan terhadap kombinasi parameter yang diatur pada *hyperparameter*. *Gradient norm* dianggap optimum ketika nilai *gradient norm* mendekati 0. Selain itu, *gradient norm* dapat menunjukkan seberapa stabil pelatihan dengan menggunakan parameter yang sudah diatur. Nilai *loss* dan *gradient norm* dihitung oleh komputer saat melatih *model chatbot*.

Evaluasi kedua yang dilakukan adalah mengevaluasi model *chatbot* menggunakan *confusion matrix*. Cara menggunakan *confusion matrix* adalah dengan menghitung akurasi yang didapatkan oleh model *chatbot* dalam menjawab pertanyaan yang berbeda. Saat melakukan evaluasi, pertanyaan yang disiapkan sebanyak 110. 5 Pertanyaan dibuat untuk 1 topik pertanyaan dan diulang sebanyak 2 kali. Selanjutnya, hasil evaluasi tersebut akan dicatatkan dalam bentuk tabel seperti Tabel 3.6. Baris horizontal pada tabel menjelaskan jawaban dari model *chatbot* dijawab sesuai dengan topik. Nomor 1 hingga nomor 10 merupakan nomor topik yang dibahas. *Outlier* memberitahukan bahwa jawaban dari model *chatbot* berada di luar topik. Kolom vertikal menjelaskan tentang pertanyaan yang sesuai topik mengikuti penomoran. Kotak warna biru menjelaskan pertanyaan yang diajukan dan jawaban yang diberikan oleh model *chatbot* sesuai. Warna hijau digunakan untuk menjumlahkan setiap pertanyaan dan topik. Hasil evaluasi ini memberikan pemahaman yang lebih mendalam terhadap kinerja model *chatbot* dalam memahami dan merespons berbagai pertanyaan dengan akurasi yang diinginkan.

Tabel 3.6 Tabel *confusion matrix*

Topik \ Jawaban	1	2	3	4	5	6	7	8	9	10	Outlier	Total
1	•	•	•	•	•	•	•	•	•	•	•	•
2	•	•	•	•	•	•	•	•	•	•	•	•
3	•	•	•	•	•	•	•	•	•	•	•	•
4	•	•	•	•	•	•	•	•	•	•	•	•
5	•	•	•	•	•	•	•	•	•	•	•	•
6	•	•	•	•	•	•	•	•	•	•	•	•
7	•	•	•	•	•	•	•	•	•	•	•	•
8	•	•	•	•	•	•	•	•	•	•	•	•
9	•	•	•	•	•	•	•	•	•	•	•	•
10	•	•	•	•	•	•	•	•	•	•	•	•
Outlier	•	•	•	•	•	•	•	•	•	•	•	•
Total	•	•	•	•	•	•	•	•	•	•	•	•

3.4.4 Analisis dan penulisan laporan

Tahapan analisis dan penulisan laporan merupakan tahapan terakhir yang dilakukan pada penelitian ini. Tahapan ini dibagi menjadi dua tahap, yang pertama adalah melakukan penghitungan akurasi dari hasil tabel *confusion matrix*, yang kedua adalah melakukan penulisan laporan. Persamaan yang digunakan untuk menghitung akurasi model diadopsi dari persamaan (2.1). Persamaan (2.1) menghitung akurasi untuk topik dengan nilai jawaban benar dan salah sedangkan *language model* pada penelitian ini dilatih dengan 10 topik. Persamaan yang digunakan untuk menghitung akurasi dengan topik jamak ada di persamaan (3.1) [15].

$$Akurasi = \frac{A}{A+B} \quad (3.1)$$

Di mana:

A = Nilai pertanyaan dan jawaban sesuai

B = Nilai pertanyaan dan jawaban tidak sesuai

V. SIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil penelitian, dapat ditarik beberapa kesimpulan sebagai berikut:

1. Parameter terbaik untuk melatih Model Chatbot Standar Operasional Prosedur Layanan Akademik di Jurusan Teknik Elektro Universitas Lampung berbasis GPT-2 adalah *learning rate* 0,002 dan *epoch* sejumlah 1000 yang menghasilkan *loss* sebesar 0,0026 dan *gradient norm* sebesar 0,01625.
2. Model Chatbot Standar Operasional Prosedur Layanan Akademik di Jurusan Teknik Elektro Universitas Lampung yang dihasilkan dari fine tuning GPT2 memiliki akurasi 20,00% dalam menangani beragam pertanyaan dengan banyak topik. Namun, kinerjanya belum baik untuk digunakan sebagai *model chatbot* standar operasional prosedur layanan akademik menggunakan *natural language processing*.
3. Model *chatbot* yang dilatih dengan baik berdasarkan evaluasi nilai *loss* dan grafik *gradient norm* tidak memiliki akurasi yang baik menurut akurasi *confusion matrix*.

5.2 Saran

Saran yang dapat digunakan untuk perkembangan penelitian selanjutnya yaitu:

1. Untuk meningkatkan akurasi, pertanyaan dengan jawaban yang sama diperbanyak kombinasi pertanyaannya.
2. Untuk meningkatkan akurasi, jumlah pertanyaan dan jawaban pada setiap topik perlu diseimbangkan.

DAFTAR PUSTAKA

- [1] “Jurusan Teknik Elektro Unila,” Jurusan Teknik Elektro Universitas Lampung, 1 November 2023. [Online]. Available: <https://ee.unila.ac.id/>. [Diakses 2 Mei 2024].
- [2] U. Lampung, “Universitas Lampung,” Universitas Lampung, 1 Desember 2023. [Online]. Available: <https://www.unila.ac.id/sejarah-universitas-lampung/>. [Diakses 1 Mei 2024].
- [3] A. Nuryanto, “Tantangan Administrasi Publik di Dunia Artificial Intelligence dan,” *Jejaring Administrasi Publik*, vol. 12, no. 2, pp. 139-147, 2020.
- [4] G. Caldarini, S. Jaf dan K. McGarry, “A Literature Survey of Recent Advances in Chatbots,” *Natural Language Interface for Smart Systems*, vol. 13, no. 1, p. 41, 2022.
- [5] N. Gupta, “A Literature Survey on Artificial Intelligence,” *International Conference on Pervasive Computing and Networking*, vol. 5, no. 19, 2017.
- [6] S. Rusell dan P. Norvig, “Agents Based on Propositional Logic,” dalam *Artificial Intelligence A Modern Approach Third Edition*, Upper Saddle River, Pearson Education Inc, 2010, p. 265.

- [7] Hugging Face Incorporated, “Hugging Face,” Hugging Face, 1 Januari 2024. [Online]. Available: <https://huggingface.co/>. [Diakses 1 Januari 2024].
- [8] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei dan I. Sutskever, “Language Models are Unsupervised Multitask Learners,” OpenAI, California, 2019.
- [9] C. Wirawan, “Huggingface,” GPT-2 Indonesia Trained by Cahya Wirawan, 18 September 2020. [Online]. Available: <https://huggingface.co/cahya/gpt2-small-indonesian-522M>. [Diakses 1 November 2023].
- [10] Python Software Foundation, “Python,” Python Software Foundation, 1 Januari 2024. [Online]. Available: <https://www.python.org/>. [Diakses 1 Januari 2024].
- [11] F. Hutter, L. Kotthoff dan J. Vanschoren, “Hyperparameter Optimization,” dalam *Automated Machine Learning*, Berlin, Springer, 2019, pp. 3-25.
- [12] F. Hutter, . L. Kotthoff dan J. Vanschoren, *Automated Machine Learning Methods, Systems, Challenges*, Berlin: Springer, 2019.
- [13] I. Goodfellow, Y. Bengio dan A. Courville, *Deep Learning*, Cambridge: MIT Press, 2016.
- [14] K. P. Murphy, *Machine Learning A Probabilistic Perspective*, Cambridge: The MIT Press, 2012, p. 57.
- [15] I. Markoulidakis, I. Rallis, I. Georgoulas, G. Kopsiaftis, A. Doulamis dan N. Doulamis, “Multiclass Confusion Matrix Reduction Method and Its

- Application on Net Promoter Score Classification Problem,” *Technologies*, vol. 9, no. 4, p. 81, 2021.
- [16] A. Huddar, C. Bysani, C. Suchak, U. D. Kolekar dan K. Upadhyaya, “Dexter the College FAQ Chatbot,” *IEEE International Conference on Convergence to Digital World*, 2020.
- [17] J. Purohit, A. Bagwe, R. Mehta, O. Mangaonkar dan E. George, “Natural Language Processing based Jaro-The Interviewing Chatbot,” dalam *Third International Conference on Computing Methodologies and Communication*, 2019.
- [18] M. M. Khan, “Development of An e-commerce Sales Chatbot,” dalam *IEEE 17th International Conference on Smart Communities: Improving Quality of Life Using ICT, IoT and AI (HONET)*, 2020.
- [19] N. P. Patel, D. R. Parikh, D. A. Patel dan R. R. Patel, “AI and Web-Based Human-Like Interactive University Chatbot (UNIBOT),” dalam *Proceedings of the Third International Conference on Electronics Communication and Aerospace Technology*, Coimbatore, 2019.
- [20] A. A. Maskutov, V. I. Zamyatovskiy, V. N. Vyuunikov dan A. V. Kutuzov, “Knowledge Base Collecting Using Natural Language Processing Algorithms,” dalam *IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIconRus)*, St. Petersburg, 2020.

- [21] A. R. M. Forkan, P. P. Jayaraman, Y.-B. Kang dan A. Morshed, "ECHO: A Tool for Empirical Evaluation Cloud Chatbots," dalam *20th IEEE/ACM International Symposium on Cluster, Cloud and Internet Computing (CCGRID)*, Melbourne, 2020.
- [22] S. Punjabi, V. Sethuram, V. Ramachandran, R. Boddu dan S. Ravi, "Chat bot Using API : Human to Machine Conversation," dalam *Global Conference for Advancement in Technology (GCAT)*, Bangalore, 2019.
- [23] H. K. K, A. K. Palakurthi, V. Putnala dan A. K. K, "Smart College Chatbot Using ML and Python," dalam *International Conference on System, Computation, Automation and Networking (ICSCAN)*, Pondicherry, 2020.
- [24] W. Badawy, A. El-Helw dan A. Youssif, "Towards Higher Customer Conversion Rate: An Interactive Chatbot Using the BEET Model," dalam *IEEE 10th International Conference on Consumer Electronics (ICCE-Berlin)*, Berlin, 2020.
- [25] M. D. Thakkar, C. U. Sanghavi, N. M. Shah dan N. Jain, "Infini - A Keyword Recognition Chatbot," dalam *International Conference on Artificial Intelligence and Smart Systems (ICAIS-2021)*, Coimbatore, 2021.
- [26] P.-C. Lin, B. Yankson, Z. Lu dan P. C. K. Hung, "Children Privacy Identification System in LINE Chatbot for Smart Toys," dalam *IEEE 12th International Conference on Cloud Computing (CLOUD)*, Milan, 2019.

- [27] Z. Yanhua, “The Application of Artificial Intelligence in Foreign Language Teaching,” dalam *International Conference on Artificial Intelligence and Education (ICAIE)*, Tianjin, 2020.
- [28] T. Chalernpol, “Information Processing and Retrieval from CSV File by Natural Language,” dalam *IEEE 3rd International Conference on Communication and Information Systems*, Bangkok, 2018.
- [29] P. Klosowski, “Deep Learning for Natural Language Processing and Language Modelling,” dalam *Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA)*, Poznan, 2018.