

**IMPLEMENTASI PENCARIAN SEMANTIK DAN *LARGE*
LANGUANGE MODELS DALAM PENERAPAN
PENCARIAN HUKUM BERDASARKAN
UNDANG-UNDANG**

SKRIPSI

M HAIKAL IRHAMNA

202000400



**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK, KOMPUTER DAN DESAIN
UNIVERSITAS NUSA PUTRA
SUKABUMI
JANUARI 2025**

**IMPLEMENTASI PENCARIAN SEMANTIK DAN *LARGE*
LANGUANGE MODELS DALAM PENERAPAN
PENCARIAN HUKUM BERDASARKAN
UNDANG-UNDANG**

SKRIPSI

*Diajukan Untuk Memenuhi Salah Satu Syarat Dalam Menempuh
Gelar Sarjana Teknik Informatika*

M HAIKAL IRHAMNA

202000400



**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK, KOMPUTER DAN DESAIN
UNIVERSITAS NUSA PUTRA
SUKABUMI
JANUARI 2025**

PERNYATAAN PENULIS

JUDUL : IMPLEMENTASI PENCARIAN SEMANTIK DAN *LARGE*
LANGUANGE MODELS DALAM PENERAPAN PENCARIAN
HUKUM BERDASARKAN UNDANG-UNDANG
NAMA : M HAIKAL IRHAMNA
NIM : 202000400

“Saya menyatakan dan bertanggungjawab dengan sebenarnya bahwa Skripsi ini adalah hasil karya saya sendiri kecuali cuplikan dan ringkasan yang masing-masing telah saya jelaskan sumbernya. Jika pada waktu selanjutnya ada pihak lain yang mengklaim bahwa Skripsi ini sebagai karyanya, yang disertai dengan bukti-bukti yang cukup, maka saya bersedia untuk dibatalkan gelar Sarjana Komputer/Sarjana Teknik saya beserta segala hak dan kewajiban yang melekat pada gelar tersebut”

Sukabumi, 18 Januari 2025



M HAIKAL IRHAMNA

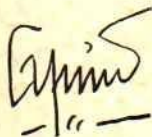
Penulis

PERSETUJUAN SKRIPSI

JUDUL : IMPLEMENTASI PENCARIAN SEMANTIK DAN *LARGE*
LANGUANGE MODELS DALAM PENERAPAN PENCARIAN
HUKUM BERDASARKAN UNDANG-UNDANG
NAMA : M HAIKAL IRHAMNA
NIM : 202000400

Skripsi ini telah diperiksa dan disetujui
Sukabumi, 18 Januari 2025

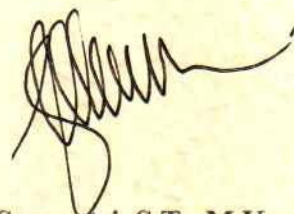
Pembimbing 1



Gina Purnama Insany, S.S.T., M.Kom

NIDN.0417077908

Pembimbing 2



Ir. Somantri, S.T., M.Kom

NIDN.0419128801

Ketua Program Studi Teknik Informatika



Ir. Somantri, S.T., M.Kom

NIDN.0419128801

PENGESAHAN SKRIPSI

JUDUL : IMPLEMENTASI PENCARIAN SEMANTIK DAN *LARGE*
LANGUANGE MODELS DALAM PENERAPAN PENCARIAN
HUKUM BERDASARKAN UNDANG-UNDANG
NAMA : M HAIKAL IRHAMNA
NIM : 202000400

Skripsi ini telah diujikan dan dipertahankan di depan Dewan Penguji pada Sidang Skripsi tanggal 3 januari 2025 Menurut pandangan kami, Skripsi ini memadai dari segi kualitas untuk tujuan penganugerahan gelar Sarjana Komputer (S.Kom).

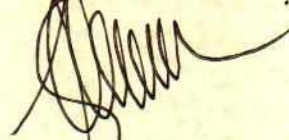
Sukabumi, 18 Januari 2025

Pembimbing 1



Gina Purnama Insany, S.S.T., M.Kom
NIDN.0417077908

Pembimbing 2



Ir. Somantri, S.T., M.Kom
NIDN.0419128801

Ketua Penguji



Alun Sujjada, S.Kom, M.T
NIDN.0718108001

Kepala Program Studi



Ir. Somantri, S.T., M.Kom
NIDN.0419128801

Plh. Dekan Fakultas Teknik, Komputer dan Desain

Ir. Paikun, S.T., M.T., IPM., ASEAN Eng
NIDN. 0402037401

ABSTRACT

Fast and accurate access to legal information is crucial for society; however, many conventional legal search systems have not yet leveraged the latest technological advancements. One potential solution is to apply semantic search technology to simplify the search for legal information based on laws. This research aims to implement semantic search in a legal information retrieval application using a WhatsApp ChatBot. The system employs the all-mpnet-base-v2 Transformer model, which maps sentences and paragraphs to a 768-dimensional vector space for clustering or semantic search. Additionally, it incorporates two comparison algorithms to measure the similarity between legal documents: Euclidean Distance and Dot Product. The data used is stored in a 768-dimensional vector form in the Qdrant database, obtained through an embedding process to generate numerical representations of legal texts. The application development method follows the Waterfall model, encompassing analysis, design, implementation, and testing phases. Testing is conducted using two approaches: Unit Testing to assess individual system functions and User Review to evaluate user experience by collecting feedback through Google Forms questionnaires. Based on the analysis of the F1-Score and threshold relationship, the Euclidean Distance method achieves an optimal F1-Score at a threshold of 0.8, whereas Dot Product achieves it at 0.58. The ROC-AUC evaluation indicates an AUC of 0.56 for Euclidean Distance (moderate performance) and 0.41 for Dot Product (less optimal performance). These comparisons demonstrate that Euclidean Distance is superior in measuring the similarity between legal documents. Testing the WhatsApp ChatBot application shows an accuracy of 90%, with a margin of error of 10%. Survey results from 51 respondents indicate a positive reception of the ChatBot: 73% of respondents find the ChatBot easy to use, 69.8% believe the ChatBot adequately understands legal terms, 71% rate the ChatBot's response speed as fast enough, and 73.8% consider the ChatBot's answers relevant. This research is expected to enhance the accuracy and efficiency of legal information retrieval, providing an innovative solution for technology-based legal information access.

Keywords: semantic search, Euclidean Distance, Dot Product, VenomBot, ChatBot, WhatsApp, vector database

ABSTRAK

Akses informasi hukum yang cepat dan akurat sangat penting bagi masyarakat, namun banyak sistem pencarian hukum konvensional yang belum memanfaatkan kemajuan teknologi terkini. Salah satu solusi yang dapat diterapkan adalah dengan menggunakan teknologi pencarian semantik untuk memudahkan pencarian informasi hukum berdasarkan undang-undang. Penelitian ini bertujuan untuk mengimplementasikan pencarian semantik dalam aplikasi pencarian hukum menggunakan *ChatBot* WhatsApp. Dalam sistem ini, pencarian hukum dilakukan dengan memanfaatkan model Transformers *all-mpnet-base-v2*, yang memetakan kalimat dan paragraf ke ruang vektor berdimensi 768 untuk pengelompokan atau pencarian semantik, serta dilengkapi dengan dua algoritma pembandingan untuk mengukur kesamaan antar dokumen hukum, yaitu Euclidean Distance dan Dot Product. Data yang digunakan disimpan dalam bentuk vektor dengan dimensi 768 di database Qdrant, yang diperoleh melalui proses embedding untuk menghasilkan representasi numerik dari teks undang-undang. Metode pengembangan aplikasi yang digunakan adalah model *Waterfall*, yang meliputi tahap analisis, desain, implementasi, dan pengujian. Pengujian dilakukan dengan dua pendekatan, yaitu Unit Testing untuk menguji fungsi-fungsi sistem secara individual dan *User Review* untuk mengevaluasi pengalaman pengguna dengan mengumpulkan umpan balik melalui kuesioner Google Form. Berdasarkan analisis hubungan F1-Score dan threshold, metode Euclidean Distance memberikan F1-Score optimal pada ambang 0.8, sedangkan Dot Product pada 0.58. Evaluasi ROC-AUC menunjukkan AUC 0.56 untuk Euclidean Distance (performa moderat) dan AUC 0.41 untuk Dot Product (performa kurang optimal). Berdasarkan perbandingan tersebut, Euclidean Distance terbukti lebih baik dalam mengukur kesamaan antar dokumen hukum. Pengujian aplikasi *ChatBot* WhatsApp menunjukkan akurasi sebesar 90%, dengan margin of error 10%. Hasil survei terhadap 51 responden menunjukkan penerimaan baik terhadap *ChatBot*. 73% responden menilai *ChatBot* cukup mudah digunakan, 69,8% responden menilai *ChatBot* cukup mampu memahami istilah hukum. 71 % menilai Kecepatan *ChatBot* dalam memberikan hasil pencarian dinilai cukup cepat, dan 73,8% menilai *ChatBot* bisa menjawab pertanyaan dengan relevan. Penelitian ini diharapkan dapat meningkatkan akurasi dan efisiensi dalam pencarian informasi hukum, serta memberikan solusi inovatif dalam akses informasi hukum berbasis teknologi.

Kata kunci: pencarian semantik, *Euclidean Distance*, *Dot Product*, *VenomBot*, *ChatBot*, WhatsApp, database vektor

KATA PENGANTAR

Alhamdulillah, puji dan syukur penulis panjatkan kehadirat Allah SWT atas segala rahmat dan karunia-Nya yang telah memberikan kekuatan serta kesabaran bagi penulis untuk menyelesaikan skripsi ini. Skripsi ini disusun sebagai salah satu syarat untuk kelulusan dan mendapatkan gelar Sarjana Komputer (S.Kom) pada Program Studi Teknik Informatika Universitas Nusa Putra Sukabumi.

Pada skripsi ini penulis mengambil judul **“IMPLEMENTASI PENCARIAN SEMANTIK DAN *LARGE LANGUAGE MODELS* DALAM PENERAPAN PENCARIAN HUKUM BERDASARKAN UNDANG-UNDANG”**. Penelitian ini bertujuan untuk menggali potensi dan efektivitas penggunaan pencarian semantik dan *large language models* dalam penerapan pencarian Undang-Undang sehingga dapat mempermudah untuk melakukan pencarian Undang-Undang.

Dalam penyusunan skripsi ini, penulis pun menemui banyak kesulitan yang disebabkan keterbatasan pengetahuan. Namun, berkat pencerahan dari Allah Yang Maha kuasa dan kemurahan hati dari setiap orang yang memberikan dukungan serta bimbingan dari berbagai sudut pemikiran, penyusunan skripsi ini dapat terselesaikan. Oleh karena itu, penulis mengucapkan terima kasih yang sedalam - dalamnya kepada semua pihak yang memberikan dukungan dalam berbagai bentuk.

Sehubungan dengan itu penulis menyampaikan penghargaan dan ucapan terima kasih yang sebesar-besarnya kepada :

1. Dr. Kurniawan, S.T., M.Si., M.M. Selaku rektor Universitas Nusa Putra Sukabumi beserta jajarannya yang telah memberikan kesempatan kepada penulis untuk menuntut ilmu di Universitas Nusa Putra.
2. Wakil Rektor I Bidang Akademik Universitas Nusa Putra Sukabumi Bapak Anggi Pradiftha Junfithrana, S.Pd., Mt beserta jajarannya yang telah memberikan kesempatan kepada penulis untuk menuntut ilmu di Universitas Nusa Putra.
3. Kepala Program Studi Teknik Informatika Bapak Somantri, S.T, M.Kom yang telah memberikan dukungan, dan pengarahan demi kelancaran penelitian ini.
4. Dosen Pembimbing I Gina Purnama Insany, S.S.T., M.Kom yang telah menyediakan waktu, tenaga, dan pikiran untuk memberikan panduan, masukan, dan dorongan dalam penyusunan skripsi ini.

5. Dosen Pembimbing II Ir. Somantri, S.T., M.Kom yang telah menyediakan waktu, tenaga, dan pikiran untuk memberikan panduan, masukan, dan dorongan dalam penyusunan skripsi ini.
6. Dosen Penguji Alun Sujjada, S.Kom, M.T Yang telah memberikan waktu dan perhatian dalam menguji skripsi ini serta memberikan masukan yang membangun.
7. Para Dosen Program Studi Teknik Informatika di Universitas Nusa Putra Sukabumi yang telah berjasa kepada penulis atas ilmu yang telah diberikan.
8. Kepada ke Orang Tua kami yang telah memberikan doa dan sangat mendukung kami dalam segala hal dan selalu mengusahakan yang terbaik.
9. Teman seperjuangan kelas TI20E angkatan 2020, terima kasih untuk kebersamaan yang telah diberikan.
10. Kepada rekan-rekan seperjuangan kami yaitu mahasiswa Program Studi Teknik Informatika yang telah menyemangati sampai selesainya skripsi ini.
11. Kepada semua pihak terkait lainnya, kami ucapkan terima kasih sebesar-besarnya.

Penulis menyadari bahwa skripsi ini masih jauh dari kesempurnaan, oleh karena itu, penulis sangat mengharapkan kritik dan saran yang mampu membangun pemikiran-pemikiran baru untuk penulisan-penulisan selanjutnya.

Sukabumi, 18 Januari 2025

M HAIKAL IRHAMNA
Penulis

**HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI
TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS**

Sebagai sivitas akademik UNIVERSITAS NUSA PUTRA, kami yang bertanda tangan di bawah ini:

NAMA : M HAIKAL IRHAMNA

NIM : 202000400

Program Studi : Teknik Informatika

Jenis karya : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Nusa Putra *Hak Bebas Royalti Noneksklusif (Non-exclusive Royalty- Free Right)* atas karya ilmiah kami yang berjudul:

“IMPLEMENTASI PENCARIAN SEMANTIK DAN *LARGE LANGUAGE MODELS* DALAM PENERAPAN PENCARIAN HUKUM BERDASARKAN UNDANG-UNDANG”. Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti *Noneksklusif* ini Universitas Nusa Putra berhak menyimpan, mengalih media/format-kan, mengelola dalam bentuk pangkalan data (database), merawat, dan memublikasikan tugas akhir kami selama tetap mencantumkan nama kami sebagai penulis/pencipta dan sebagai pemilik Hak Cipta. Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Universitas Nusa Putra

Pada tanggal : 18 Januari 2025

Yang Menyatakan:



M HAIKAL IRHAMNA
Penulis

DAFTAR ISI

PERNYATAAN PENULIS.....	iii
PERSETUJUAN SKRIPSI.....	iv
PENGESAHAN SKRIPSI	v
HALAMAN PERSEMBAHAN.....	vi
ABSTRACT.....	vii
ABSTRAK.....	viii
KATA PENGANTAR	ix
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI.....	xi
DAFTAR ISI.....	xii
DAFTAR TABEL.....	xv
DAFTAR GAMBAR.....	xvi
DAFTAR LAMPIRAN	xviii
BAB I.....	1
PENDAHULUAN	1
11.1 Latar Belakang.....	1
11.2 Rumusan Masalah.....	4
11.3 Batasan Masalah	4
11.4 Tujuan Penelitian	5
11.5 Manfaat Penelitian	5
11.6 Sistematika Penulisan	7
BAB II.....	8
TINJAUAN PUSTAKA	8
2.1 Penelitian Terkait.....	8
2.2 Pencarian Semantik.....	11
2.3 Metode <i>Euclidean Distance</i>	12
2.4 <i>Dot Product</i>	14

2.5	<i>Large Language Models (LLM)</i>	15
2.6	Transformers model all-mpnet-base-v2.....	16
2.7	Google Api Translate	16
2.8	Venom bot	17
2.9	Python.....	17
2.10	Node.js	18
2.11	Framework Flask	19
2.12	Model Proses <i>Waterfall</i>	20
2.13	<i>Blackbox testing</i>	21
2.14	<i>Unit Testing Dan User Review</i>	22
2.15	Integrasi Teknologi dalam Hukum	22
2.16	Unified Modeling Language (UML).....	24
2.17	Use Case Diagram.....	25
2.18	Sequence Diagram	26
2.19	Activity Diagram.....	27
2.20	Class Diagram	28
2.21	<i>Deployment Diagram</i>	29
2.22	Kerangka Pemikiran.....	30
BAB III		34
METODE PENELITIAN		34
3.1	Tahapan Penelitian.....	34
3.2	Mengidentifikasi Masalah	35
3.3	Perancangan Sistem	37
3.4	Metode Pengumpulan Data.....	40
3.4.1	Studi Pustaka.....	40
3.4.2	Kuesioner	41
3.4.3	Data Mining	41

3.5	<i>Data Preprocessing</i>	44
3.6	Mode Selection	47
3.6.1	Euclidean Distance	47
3.6.2	Dot Product	51
3.6.3	Pencarian Semantik	52
3.7	Model Testing.....	53
3.8	<i>Deployment</i>	53
3.9	<i>Metode Testing</i>	54
3.10	Maintenance	54
3.11	Hasil Dan kesimpulan	54
BAB IV	HASIL DAN PEMBAHASAN.....	56
4.1	Data Mining	56
4.2	Data Preprocessing.....	56
4.3	Implementasi Algoritma	62
4.4	<i>Large Language Model</i>	64
4.5	Model Testing.....	65
4.6	<i>Deployment</i>	67
4.7	Metode Testing	69
4.7.1	Unit Testing	71
4.7.2	<i>User Review</i>	73
BAB V	PENUTUP	80
5.1	Kesimpulan	80
5.2	Saran	81
LAMPIRAN		87

DAFTAR TABEL

Tabel 2. 2 Penelitian Terkait	8
Tabel 2. 3 UML <i>view</i> dan <i>Diagram</i> [46].....	24
Tabel 2. 4 Macam-macam Notasi <i>Use Case Diagram</i> [46]	25
Tabel 2. 5 Simbol <i>Use Case Diagram</i> [46].....	26
Tabel 2. 6 Simbol <i>Sequence Diagram</i> [46].....	27
Tabel 2. 7 Simbol <i>Activity Diagram</i> [46].....	27
Tabel 2. 8 Simbol <i>Class Diagram</i> [46].....	28
Tabel 2. 9 Simbol <i>Deployment Diagram</i> [46].....	29



DAFTAR GAMBAR

Gambar 2. 1 Ilustrasi Pencarian Semantik	12
Gambar 2. 2 Metode <i>Waterfall</i> [42]	21
Gambar 2. 3 Kerangka Pemikiran	32
Gambar 3. 1 Tahapan Penelitian	34
Gambar 3. 2 Use Case Sistem	37
Gambar 3. 3 <i>Activity Send Request</i>	38
Gambar 3. 4 <i>Backend Processing</i>	39
Gambar 3. 5 Data mining	42
Gambar 3. 6 Algoritma <i>Euclidean Distance</i>	48
Gambar 4. 1 Pdf Hasil Download.....	56
Gambar 4. 2 Tahapan Data Preprosesing	57
Gambar 4. 3 Data File Json	57
Gambar 4. 4 Kode Menggunakan API google	58
Gambar 4. 5 kode embedding word	59
Gambar 4. 6 Konfigurasi <i>Vector 1</i>	59
Gambar 4. 7 konfigurasi <i>vector 2</i>	60
Gambar 4. 8 data <i>Vector</i>	61
Gambar 4. 9 Contoh data yang disimpan di qdran 1	61
Gambar 4. 10 Contoh data yang disimpan di qdran 1	62
Gambar 4. 11 hasil yang paling baik.....	63
Gambar 4. 12 hasil ke dua paling baik	63
Gambar 4. 13 hasil pencarian tidak relevan	64
Gambar 4. 14 <i>Prompt LLM to get summary</i>	64
Gambar 4. 15 F1-Score vs <i>Threshold Euclidean Distance</i>	65
Gambar 4. 16 F1-Score vs <i>Threshold Dot Product</i>	66
Gambar 4. 17 ROC-AUC Score <i>Euclidean Distance</i>	66
Gambar 4. 18 ROC-AUC Score <i>Dot Product</i>	67
Gambar 4. 19 <i>Arsitektur system ChatBot</i>	68
Gambar 4. 20 <i>ChatBot Whatsapp</i>	69
Gambar 4. 22 Garfik batang kemudahan penggunaan <i>ChatBot</i>	74
Gambar 4. 23 Grafik Mengenai pemahaman istilah.....	75

Gambar 4. 24 Mengenai kecepatan <i>ChatBot</i> menjawab pertanyaan	76
Gambar 4. 25 Grafik Mengenai relevansi jawaban <i>ChatBot</i>	77



DAFTAR LAMPIRAN

Lampiran 1 Kode untuk mengunduh berkas pdf 1	87
Lampiran 2 Kode untuk mengunduh berkas pdf 2	87
Lampiran 3 Kode untuk mengekstraksi teks dari berkas pdf 1	88
Lampiran 4 Kode untuk mengekstraksi teks dari berkas pdf 2	88
Lampiran 5 Kode untuk mengekstraksi teks dari berkas pdf 3	89
Lampiran 6 Kode untuk mengekstraksi teks dari berkas pdf 4	89
Lampiran 7 Kode untuk menyimpan data ke database 1	90
Lampiran 8 Kode untuk menyimpan data ke database 2	90
Lampiran 9 Kode untuk menyimpan data ke database 3	91
Lampiran 10 Kode untuk menyimpan data ke database 4	91
Lampiran 11 Kode untuk mengambil data dari database 1	92
Lampiran 12 Kode untuk mengambil data dari database 2	92
Lampiran 13 Kode untuk mengambil data dari database 3	93
Lampiran 14 Kode untuk menyimpan data ke qdrant 1	93
Lampiran 15 Kode untuk menyimpan data ke qdrant 2	94
Lampiran 16 Kode untuk menjalankan fungsi utama pengolahan data 1	94
Lampiran 17 Kode untuk menjalankan fungsi utama pengolahan data 2	95
Lampiran 18 Kode untuk menjalankan fungsi utama pengolahan data 3	95
Lampiran 19 Kode untuk fungsi terjemah	96
Lampiran 20 kode untuk fungsi open ai	96
Lampiran 21 kode untuk fungsi <i>embedding</i>	96
Lampiran 22 Kode untuk mengambil data dari <i>database</i> 1	97
Lampiran 23 Kode untuk mengambil data dari <i>database</i> 2	97
Lampiran 24 Kode untuk melakukan pencarian semantik 1	98
Lampiran 25 Kode untuk melakukan pencarian semantik 2	98
Lampiran 26 Kode untuk melakukan pencarian semantik 3	99
Lampiran 27 Kode untuk melakukan pencarian semantik 4	99
Lampiran 28 Kode untuk menjalankan flask	99
Lampiran 29 Kode Chat bot 1	100
Lampiran 30 Kode Chat bot 2	100
Lampiran 31 Kode Chat bot 3	101

Lampiran 32 Kode Chat bot 4.....	101
Lampiran 33 Kode untuk melakukan <i>unit testing</i> 1	102
Lampiran 34 Kode untuk melakukan <i>unit testing</i> 2	102
Lampiran 35 Kode untuk melakukan <i>unit testing</i> 3	103
Lampiran 36 Kode untuk melakukan <i>unit testing</i> 4	103
Lampiran 37 Kode untuk melakukan <i>unit testing</i> 5	104
Lampiran 38 Hasil <i>unit testing</i>	104
Lampiran 39 Hasil Pencarian 1	105
Lampiran 40 Hasil Pencarian 2	105
Lampiran 41 Hasil Pencarian 3	106
Lampiran 42 Hasil Pencarian 4	106
Lampiran 43 Hasil Pencarian 5	107
Lampiran 44 Hasil Pencarian 6	107
Lampiran 45 Hasil Pencarian 7	108
Lampiran 46 Hasil Pencarian 8	108
Lampiran 47 Hasil Pencarian 9	109
Lampiran 48 Hasil Pencarian 10	109



BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi informasi telah membawa perubahan signifikan dalam berbagai aspek kehidupan, termasuk dalam bidang hukum[1]. Pencarian dan penelusuran informasi hukum yang efisien dan akurat menjadi kebutuhan mendasar bagi praktisi hukum, akademisi, maupun masyarakat umum. Namun, sistem pencarian konvensional yang mengandalkan kata kunci (*keyword-based search*) seringkali menghadapi kendala dalam memahami konteks dan makna yang lebih dalam dari *query* pengguna [2].

Di era digital saat ini, platform perpesanan seperti WhatsApp telah menjadi saluran komunikasi utama bagi mayoritas masyarakat Indonesia, dengan penetrasi pengguna mencapai 84.8% dari total pengguna internet [3]. Integrasi layanan pencarian hukum melalui *ChatBot* WhatsApp menjadi solusi strategis untuk memperluas akses masyarakat terhadap informasi hukum[4]. penggunaan *ChatBot* berbasis WhatsApp dapat meningkatkan efisiensi dan aksesibilitas layanan publik dibandingkan dengan platform konvensional.

Undang-undang sebagai sumber hukum formal memiliki struktur bahasa yang kompleks dan saling terkait antara satu pasal dengan pasal lainnya[5]. Kompleksitas ini menyebabkan pencarian informasi hukum menjadi tantangan tersendiri, terutama ketika pengguna ingin menemukan interpretasi atau penerapan hukum yang relevan dengan kasus tertentu. Penelitian yang dilakukan oleh (Sastriani et al., 2020)[6], menunjukkan bahwa sistem pencarian tradisional yang hanya mencocokkan kata kunci seringkali menghasilkan hasil yang kurang relevan dan membutuhkan waktu yang lebih lama dalam proses penelusuran.

Kemajuan dalam bidang kecerdasan buatan, khususnya dalam teknologi pemrosesan bahasa alami (*Natural Language Processing*), telah memunculkan dua pendekatan yang menjanjikan: pencarian semantik dan *Large Language Models* (LLM). Menurut (Herwanza et al., 2024)[7], pencarian semantik memungkinkan sistem untuk memahami makna dan konteks dari *query* pengguna, tidak hanya

sebatas pencocokan kata. Sementara itu, LLM seperti GPT (*Generative Pre-trained Transformer*) memiliki kemampuan untuk memahami dan menganalisis teks dalam skala besar dengan pemahaman kontekstual yang lebih baik[8].

Integrasi *ChatBot* WhatsApp dengan teknologi LLM dan pencarian semantik membuka peluang baru dalam *demokratisasi* akses informasi hukum. Studi yang dilakukan oleh (Lubis et al., 2024)[9], menunjukkan bahwa implementasi *ChatBot* berbasis LLM dapat meningkatkan tingkat kepuasan pengguna hingga 78% dalam mengakses informasi hukum. Sistem ini memungkinkan pengguna untuk mengajukan pertanyaan dalam bahasa natural dan mendapatkan respons yang lebih kontekstual dan mudah dipahami.

Penelitian yang dilakukan oleh (Lai et al., 2024)[10], menunjukkan bahwa penggabungan kedua teknologi ini berpotensi memberikan solusi yang lebih efektif dalam pencarian dan analisis informasi hukum. Pencarian semantik dapat meningkatkan presisi hasil pencarian dengan memahami maksud pengguna, sedangkan LLM dapat membantu dalam menginterpretasikan dan menjelaskan konteks hukum yang relevan [11]. Namun, implementasi teknologi ini dalam konteks sistem hukum Indonesia masih memerlukan kajian mendalam, terutama terkait dengan akurasi, efektivitas, dan kesesuaiannya dengan kerangka hukum yang berlaku.

ChatBot berbasis teknologi LLM yang diimplementasikan pada platform Whatsapp juga dapat digunakan untuk meningkatkan interaksi pengguna dengan sistem pencarian hukum. *ChatBot* ini dapat memberikan respons otomatis dan relevan terhadap pertanyaan hukum yang diajukan oleh pengguna, membantu mereka menemukan informasi yang dibutuhkan dengan lebih cepat dan efisien. Penggunaan *ChatBot* dalam konteks ini memungkinkan pengguna untuk berinteraksi dengan sistem menggunakan bahasa alami, yang dapat meningkatkan kenyamanan dan efisiensi pencarian.

Undang-Undang (UU) adalah Peraturan Perundang-undangan yang dibentuk oleh Dewan Perwakilan Rakyat dengan persetujuan bersama Presiden, peraturan perundang-undangan yang diakui keberadaannya mempunyai kekuatan hukum mengikat[12]. dan penerapan hukum dilakukan dengan benar dan adil. Penggunaan LLM dan pencarian semantik dapat membantu dalam memahami dan menafsirkan

ketentuan-ketentuan dalam UU, serta menyediakan akses yang lebih mudah dan cepat terhadap informasi yang relevan.

Untuk meningkatkan akurasi pencarian, metode *Euclidean* dapat digunakan dalam pengukuran kesamaan antara dokumen hukum. Metode *Euclidean* memungkinkan pengukuran jarak antara representasi *vector* dari dokumen, sehingga dapat membantu dalam mengidentifikasi dokumen yang paling relevan berdasarkan konten dan konteks yang dicari[13]. Penggunaan metode ini dapat meningkatkan presisi dan relevansi hasil pencarian, memberikan hasil yang lebih akurat dan kontekstual.

ChatBot dengan fitur pencarian semantik ini sangat penting, terutama bagi generasi muda dan masyarakat umum yang tidak memiliki latar belakang hukum yang kuat. Generasi muda saat ini sangat akrab dengan teknologi dan internet, namun seringkali kesulitan dalam memahami dan mencari informasi hukum yang relevan. Dengan adanya *ChatBot* yang memanfaatkan pencarian semantik di Whatsapp, mereka dapat lebih mudah mengakses dan memahami dokumen hukum, termasuk UU.

Selain itu, dalam banyak kasus hukum yang terjadi di masyarakat, informasi yang akurat dan relevan sering kali sulit ditemukan. Pencarian hukum tradisional yang hanya mengandalkan kata kunci dapat menyebabkan ketidakakuratan dalam hasil pencarian dan menghabiskan banyak waktu. *ChatBot* dengan pencarian semantik dapat membantu dalam menyediakan informasi yang lebih cepat dan tepat, sehingga dapat membantu dalam penyelesaian kasus-kasus hukum dengan lebih efisien.

Penelitian ini bertujuan untuk mengkaji implementasi pencarian semantik dan LLM dalam penerapan pencarian hukum berdasarkan Undang-undang. Melalui penelitian ini, diharapkan dapat ditemukan metode yang lebih efektif dan efisien dalam mencari dan menafsirkan informasi hukum yang kompleks, sehingga dapat memberikan kontribusi positif terhadap praktik hukum di Indonesia. Dengan adanya penelitian ini, diharapkan juga dapat memberikan wawasan baru mengenai potensi dan tantangan dalam penggunaan teknologi canggih di bidang hukum, serta memberikan rekomendasi yang dapat diimplementasikan oleh instansi pemerintah, lembaga hukum, dan pihak-pihak terkait lainnya.

Dengan adanya penelitian ini, diharapkan juga dapat memberikan wawasan baru mengenai potensi dan tantangan dalam penggunaan teknologi canggih di bidang hukum, serta memberikan rekomendasi yang dapat diimplementasikan oleh instansi pemerintah, lembaga hukum, dan pihak-pihak terkait lainnya.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dipaparkan diatas, maka permasalahan yang didapat adalah sebagai berikut :

1. Bagaimana implementasi pencarian semantik dan *Large Language Models* (LLM) dapat meningkatkan akurasi dan relevansi hasil pencarian hukum berdasarkan Undang-Undang(UU) Indonesia?
2. Bagaimana pencarian semantik dan LLM dapat mempermudah akses informasi hukum bagi generasi muda dan masyarakat umum yang tidak memiliki latar belakang hukum?
3. Bagaimana efektivitas pencarian semantik dengan menggunakan bahasa inggris dan LLM dalam membantu penyelesaian kasus-kasus hukum di Indonesia?
4. Bagaimana perbandingan performa metode Euclidean Distance dan Dot Product dalam pencarian semantik, berdasarkan evaluasi F1-Score dan ROC-AUC serta akurasi dari *ChatBot* whatsapp?
5. Sejauh mana penerimaan pengguna terhadap *ChatBot* sebagai alat pencarian informasi hukum, dan apa saja aspek yang perlu diperbaiki untuk meningkatkan performa *ChatBot* tersebut?

1.3 Batasan Masalah

Agar pembahasan penelitian ini berfokus pada ruang lingkup masalah yang diinginkan, maka ada batasan masalah yang akan dibatasi sebagai berikut :

1. Penelitian ini terbatas pada implementasi pencarian semantik dan Model Bahasa Besar (LLM) dalam konteks pencarian hukum berdasarkan Undang-Undang(UU) Indonesia.
2. Data yang digunakan dalam penelitian ini terbatas pada dokumen hukum yang relevan dengan UU Indonesia dan kasus-kasus hukum terkait.

3. Pengujian dan evaluasi dilakukan pada prototipe *ChatBot* WhatsApp untuk pencarian hukum yang dikembangkan selama penelitian, tidak mencakup implementasi skala penuh pada sistem hukum nasional.
4. Dataset yang tersedia adalah undang-undang KUHP yang di dapat dari BPHN.JDIH.

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah diatas, maka penelitian ini bertujuan untuk :

1. Mengimplementasikan pencarian semantik dan Model Bahasa Besar (LLM) untuk dapat meningkatkan akurasi dan relevansi hasil pencarian hukum berdasarkan Undang-Undang.
2. Menerapkan pencarian semantik dan LLM untuk dapat mempermudah akses informasi hukum bagi generasi muda dan masyarakat umum yang tidak memiliki latar belakang hukum.
3. Mengetahui efektivitas pencarian semantik dan LLM dalam membantu penyelesaian kasus-kasus hukum di Indonesia.
4. Membandingkan kinerja metode Euclidean Distance dan Dot Product dalam pencarian semantik, dengan menggunakan metrik F1-Score dan ROC-AUC sebagai parameter evaluasi serta mengetahui akurasi dari *ChatBot* whatsapp yang di kembangkan.
5. Mengukur penerimaan pengguna terhadap penggunaan *ChatBot* dalam mencari informasi hukum, serta mengidentifikasi aspek-aspek yang perlu diperbaiki berdasarkan survei responden.

1.5 Manfaat Penelitian

Manfaat dari penelitian ini adalah:

1. Untuk Mahasiswa
 - a. Peningkatan Pemahaman Hukum: Penelitian ini memberikan pemahaman yang lebih mendalam tentang bagaimana teknologi seperti pencarian semantik dan LLM dapat diterapkan dalam bidang hukum, membantu mahasiswa memahami konsep-konsep canggih dalam teknologi informasi dan hukum.

- b. Kemudahan Akses Informasi: Dengan adanya *ChatBot* di WhatsApp pencarian hukum yang lebih cerdas, mahasiswa dapat dengan mudah mengakses informasi hukum yang relevan dan akurat untuk mendukung studi dan penelitian mereka.
 - c. Keterampilan Teknologi: Mahasiswa yang terlibat dalam penelitian ini akan mendapatkan keterampilan praktis dalam pengembangan dan implementasi teknologi pencarian semantik dan LLM, yang sangat berharga dalam karir mereka di bidang teknologi informasi dan hukum.
2. Untuk Universitas Nusa Putra
- a. Peningkatan Reputasi Akademik: Penelitian ini dapat meningkatkan reputasi Universitas Nusa Putra sebagai institusi yang inovatif dan berkontribusi dalam pengembangan teknologi canggih di bidang hukum.
 - b. Pengembangan Kurikulum: Hasil penelitian dapat digunakan sebagai bahan ajar dan referensi untuk pengembangan kurikulum di jurusan terkait, seperti teknologi informasi, hukum, dan ilmu komputer.
 - c. Kerjasama dan Kolaborasi: Penelitian ini membuka peluang bagi universitas untuk menjalin kerjasama dengan institusi hukum dan teknologi, baik di tingkat nasional maupun internasional.
3. Untuk Masyarakat
- a. Akses Informasi Hukum yang Lebih Baik: Masyarakat umum, termasuk generasi muda, dapat dengan mudah mengakses dan memahami informasi hukum yang relevan dan akurat, tanpa memerlukan latar belakang hukum yang kuat.
 - b. Penyelesaian Kasus Hukum yang Lebih Efisien: Dengan pencarian yang lebih akurat dan relevan, masyarakat dapat menemukan informasi yang mereka butuhkan untuk menyelesaikan kasus hukum dengan lebih efisien.
 - c. Peningkatan Literasi Hukum: Penelitian ini dapat meningkatkan literasi hukum di kalangan masyarakat, membantu mereka memahami hak dan kewajiban mereka sesuai dengan Undang-undang, sehingga dapat berpartisipasi secara lebih aktif dan informatif dalam proses hukum.

1.6 Sistematika Penulisan

Sistematika penulisan skripsi ini disusun untuk memberikan gambaran secara umum tentang penulisan skripsi yang dikerjakan. Adapun sistematika skripsi ini adalah sebagai berikut :

BAB I: Pendahuluan

Pada bab ini dijelaskan tentang latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian dan sistematika penulisan

BAB II: Tinjauan Pustaka

Pada bab ini membahas tentang landasan umum mengenai konsep dasar dan teori yang berkaitan dengan penelitian diantaranya mengenai literatur-literatur terkait tentang metode pencarian semantik, *Large Language Models*, *ChatBot*, Undang-Undang, serta penelitian terdahulu yang relevan.

BAB III: Metodologi Penelitian

Pada bab ini mencakup berbagai tahapan penelitian, meliputi tahapan pengumpulan data, analisis, perancangan, dan pengembangan sistem. Dibahas juga teknik dan alat penelitian yang digunakan sepanjang tahapan tersebut, serta lokasi subjek penelitian, data/informasi yang diperlukan, dan waktu yang dialokasikan untuk setiap tahapan.

BAB IV: Hasil Dan Pembahasan

Pada bab ini berisi tentang hasil dari penelitian yang dilakukan, pembahasan dari penelitian yang dilakukan dan pembahasan dari implementasi pencarian semantik dan *large language models* Pda *ChatBot* WhatsApp, serta pengujian dari penelitian yang telah dikerjakan.

BAB V : Penutup

Pada bab ini berisi kesimpulan dari semua tahapan penelitian yang dilakukan dan saran – saran untuk pengembangan penelitian selanjutnya.

BAB V PENUTUP

1.1 Kesimpulan

Berdasarkan penelitian yang dilakukan dengan judul *Implementasi Pencarian Semantik dan Large Language Models dalam Penerapan Pencarian Hukum Berdasarkan Undang-Undang* menggunakan algoritma *Euclidean Distance*, dapat disimpulkan sebagai berikut:

1. Peningkatan Akurasi dan Relevansi, Implementasi pencarian semantik yang didukung oleh Large Language Models (LLM) mampu meningkatkan akurasi dan relevansi hasil pencarian hukum berdasarkan Undang-Undang Indonesia. Hal ini disebabkan oleh kemampuan LLM dalam memahami konteks dan makna semantik dari dokumen hukum, sehingga hasil pencarian lebih sesuai dengan kebutuhan pengguna.
2. Kemudahan Akses Informasi Hukum, Penerapan pencarian semantik dan LLM memberikan kemudahan bagi generasi muda dan masyarakat umum yang tidak memiliki latar belakang hukum dalam mengakses informasi hukum. Antarmuka pencarian yang dirancang sederhana dan kemampuan sistem dalam menangkap pertanyaan berbasis bahasa alami membuat akses informasi hukum menjadi lebih inklusif.
3. Efektivitas dalam Penyelesaian Kasus Hukum Sistem pencarian berbasis LLM dan semantik efektif dalam membantu penyelesaian kasus hukum. Hal ini ditunjukkan oleh kemampuannya untuk memberikan referensi Undang-Undang yang relevan secara cepat dan akurat, sehingga mempercepat proses analisis hukum bagi para praktisi hukum maupun masyarakat.
4. Berdasarkan analisis hubungan F1-Score dan threshold, metode Euclidean Distance memberikan F1-Score optimal pada ambang 0.8, sedangkan Dot Product pada 0.58. Evaluasi ROC-AUC menunjukkan AUC 0.56 untuk Euclidean Distance (performa moderat) dan AUC 0.41 untuk Dot Product (performa kurang optimal). Berdasarkan perbandingan Euclidean Distance lebih baik. Dan berdasarkan pengujian yang dilakukan akurasi aplikasi *ChatBot WhatsApp* adalah 90%, dengan margin of error 10%.

5. Berdasarkan survei yang dilakukan terhadap 51 responden mengenai penggunaan *ChatBot* untuk mencari informasi hukum, hasil yang diperoleh menunjukkan adanya penerimaan yang baik terhadap *ChatBot*, meskipun masih ada ruang untuk perbaikan. Pada aspek kemudahan penggunaan, sebagian besar responden (41,2%) menilai *ChatBot* cukup mudah digunakan (skala 3). Dalam hal pemahaman terhadap istilah hukum, 47,1% responden merasa *ChatBot* cukup mampu memahami pertanyaan (skala 3). Kecepatan *ChatBot* dalam memberikan hasil pencarian hukum dinilai cukup cepat oleh 52,9% responden (skala 3). Terakhir, terkait relevansi hasil pencarian, 43,1% responden menilai hasilnya cukup relevan (skala 4). Hasil ini menunjukkan adanya potensi untuk meningkatkan beberapa aspek *ChatBot*.

1.2 Saran

Berdasarkan hasil penelitian, berikut adalah beberapa saran untuk pengembangan dan implementasi lebih lanjut:

1. Peningkatan Basis Data Hukum, Sistem perlu diintegrasikan dengan basis data hukum yang lebih luas dan terkini, mencakup peraturan pemerintah, peraturan daerah, serta putusan pengadilan untuk meningkatkan cakupan pencarian dan relevansi hasil.
2. Peningkatan Teknologi Algoritma Algoritma *Euclidean Distance* dapat dikombinasikan dengan algoritma lain seperti *cosine similarity* atau teknologi berbasis pembelajaran mendalam (*deep learning*) untuk meningkatkan keakuratan dalam penentuan kesamaan dokumen.
3. Pelatihan Model Berbasis Konteks Lokal, Mengembangkan model LLM yang lebih spesifik terhadap konteks hukum Indonesia, dengan melatih model menggunakan data dari dokumen hukum lokal, agar lebih adaptif terhadap terminologi hukum yang unik di Indonesia.
4. Pengembangan Antarmuka Pengguna, Sistem dapat dilengkapi dengan antarmuka visual yang lebih ramah pengguna serta fitur tambahan seperti rekomendasi otomatis, penjelasan pasal, dan navigasi yang lebih intuitif untuk meningkatkan pengalaman pengguna.

5. Sosialisasi dan Edukasi, Sosialisasi penting dilakukan untuk memperkenalkan sistem ini kepada masyarakat umum, khususnya generasi muda, melalui pelatihan atau seminar terkait pemanfaatan teknologi pencarian hukum berbasis LLM.
6. Evaluasi Berkelanjutan, Dilakukan evaluasi berkala terhadap kinerja sistem dengan melibatkan pengguna dari berbagai latar belakang, baik hukum maupun non-hukum, untuk memastikan sistem tetap relevan dan bermanfaat.

Dengan implementasi dan pengembangan yang berkesinambungan, sistem pencarian semantik berbasis LLM dapat menjadi solusi inovatif dalam mempermudah akses dan pemanfaatan informasi hukum di Indonesia.



DAFTAR PUSTAKA

- [1] R. Susskind, *Online courts and the future of justice*. Oxford University Press, 2019.
- [2] K. D. Ashley, *Artificial intelligence and legal analytics: new tools for law practice in the digital age*. Cambridge University Press, 2017.
- [3] W. A. Social, “Digital 2019: Indonesia, We Are Social & Hootsuite.” London, 2019.
- [4] B. H. untuk Akses, “Keterbukaan Informasi Bantuan Hukum untuk Akses Keadilan yang Lebih Luas,” 2023.
- [5] M. H. Muhtar *et al.*, “Konsep Hukum Indonesia,” *Glob. Eksek. Teknol.*, vol. 35, 2023.
- [6] R. Sastriani, Z. K. A. Baizal, and D. S. Kusumo, “Ontology-based semantic search on tourism information search system,” *Indones. J. Comput.*, vol. 5, no. 1, pp. 103–114, 2020.
- [7] N. A. M. Herwanza, N. S. Harahap, F. Yanto, and F. Insani, “Penerapan Langchain Retriever dengan Model Chat Openai dalam Pengembangan Sistem ChatBot Hadis Berbasis Telegram,” *JTIM J. Teknol. Inf. dan Multimed.*, vol. 6, no. 1, pp. 70–83, 2024.
- [8] F. Sulianta, *Chat GPT-Memberdayakan Large Language Model untuk Berbagai Kebutuhan*. Feri Sulianta, 2024.
- [9] A. T. U. B. R. Lubis, N. S. Harahap, S. Agustian, M. Irsyad, and I. Afrianty, “Question Answering System pada ChatBot Telegram Menggunakan Large Language Models (LLM) dan Langchain (Studi Kasus UU Kesehatan): Question Answering System on Telegram ChatBot Using Large Language Models (LLM) and Langchain (Case Study: Health Law),” *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 4, no. 3, pp. 955–964, 2024.
- [10] J. Lai, W. Gan, J. Wu, Z. Qi, and S. Y. Philip, “Large language models in law: A survey,” *AI Open*, 2024.
- [11] S. Ma, C. Chen, Q. Chu, and J. Mao, “Leveraging Large Language Models for Relevance Judgments in Legal Case Retrieval,” *arXiv Prepr. arXiv2403.18405*, 2024.
- [12] B. VERDINAND, “Fungsi Legislasi Dewan Perwakilan Rakyat Republik Indonesia Dalam Undang-Undang Republik Indonesia Nomor 15 Tahun 2019 Tentang Perubahan Atas Undang-Undang Nomor 12 Tahun 2011 Tentang Pembentukan Peraturan Perundang-Undangan.” Universitas Jambi, 2022.
- [13] S. Rimadias, “Optimalisasi Artificial Intelligence (AI) Pada Platform SINTA,” *Sinta DIkti*, 2024.
- [14] Y. Han, C. Liu, and P. Wang, “A Comprehensive Survey on Vector Database: Storage and Retrieval Technique, Challenge,” *arXiv Prepr. arXiv2310.11703*, 2023.

- [15] D. Schumacher, "V3CTRON| Data Retrieval & Access System For Flexible Semantic Search & Retrieval Of Proprietary Document Collections Using Natural Language Queries.," *Available SSRN*, 2023.
- [16] N. Oh, G.-S. Choi, and W. Y. Lee, "ChatGPT goes to the operating room: evaluating GPT-4 performance and its potential in surgical education and training in the era of large language models," *Ann. Surg. Treat. Res.*, vol. 104, no. 5, p. 269, 2023.
- [17] M. A. Izza, A. Jazuli, and M. Nurkamid, "Implementasi Teknologi Semantik Web Untuk Pencarian Koleksi Perpustakaan Universitas Muria Kudus," *J. Dialekt. Inform.*, vol. 2, no. 2, pp. 56–62, 2022.
- [18] A. Andri, "IMPLEMENTASI TEKNOLOGI WEB SEMANTIK DALAM APLIKASI PENCARIAN KATALOG ONLINE PERPUSTAKAAN UNIVERSITAS BINA DARMA," *IMPLEMENTASI Teknol. WEB Semant. DALAM Apl. PENCARIAN KATALOG ONLINE Perpust. Univ. BINA DARMA*, 2022.
- [19] V. Setlur, A. Kanyuka, and A. Srinivasan, "Olio: A Semantic Search Interface for Data Repositories," in *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*, 2023, pp. 1–16.
- [20] E. H. Mohamed and E. M. Shokry, "QSST: A Quranic Semantic Search Tool based on word embedding," *J. King Saud Univ. Inf. Sci.*, vol. 34, no. 3, pp. 934–945, 2022.
- [21] H. S. Fardini and R. Septiriana, "PERBANDINGAN NILAI SIMILARITY EUCLIDEAN DISTANCE DAN COSINE SIMILARITY UNTUK DIAGNOSIS PENYAKIT GIGI DAN MULUT MENGGUNAKAN PENALARAN BERBASIS KASUS," *Sci. J. Ilm. Sains dan Teknol.*, vol. 2, no. 5, pp. 103–117, 2024.
- [22] A. T. Wijaya and S. Subandi, "PENERAPAN METODE CLUSTERING DENGAN ALGORITMA K-MEANS PADA SISTEM PENDETEKSI PENCUCIAN UANG PERBANKAN BERBASIS WEB," in *Prosiding Seminar Nasional Mahasiswa Fakultas Teknologi Informasi (SENAFTI)*, 2024, pp. 398–406.
- [23] A. Radford *et al.*, "Learning transferable visual models from natural language supervision," in *International conference on machine learning*, PMLR, 2021, pp. 8748–8763.
- [24] N. G. J. D. M. Joshi, D. C. O. L. M. Lewis, L. Z. V. S. Y. Liu, M. Ott, and V. Stoyanov, "Roberta: A robustly optimized bert pretraining approach," in *Submitted to International Conference on Learning Representations*. <https://openreview.net/forum>, 2020.
- [25] S. Ruder, M. E. Peters, S. Swayamdipta, and T. Wolf, "Transfer learning in natural language processing," in *Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: Tutorials*, 2019, pp. 15–18.

- [26] J. Devlin, "Bert: Pre-training of deep bidirectional transformers for language understanding," *arXiv Prepr. arXiv1810.04805*, 2018.
- [27] W. X. Zhao *et al.*, "A survey of large language models," *arXiv Prepr. arXiv2303.18223*, 2023.
- [28] H. Naveed *et al.*, "A comprehensive overview of large language models," *arXiv Prepr. arXiv2307.06435*, 2023.
- [29] C. Galli, N. Donos, and E. Calciolari, "Performance of 4 Pre-Trained Sentence Transformer Models in the Semantic Query of a Systematic Review Dataset on Peri-Implantitis," *Information*, vol. 15, no. 2, p. 68, 2024.
- [30] N. Andindita, "Aplikasi Translator Messenger Berbasis Java dan Google Translate API," *Univ. Negeri Yogyakarta. http://eprints.uny.ac.id/10380/1/Journal.pdf (diakses tanggal 1 Maret 2014)*, 2013.
- [31] F. B. Hawari, "Pemanfaatan Google Cloud Speech Api Dan Alchemy Api Pada Aplikasi Perekam Wawancara Jurnalis Berbasis Android." Universitas Komputer Indonesia, 2018.
- [32] R. Suryana, M. Aryanto, R. Kurniawan, K. S. G. P. Satmata, A. S. Yulianti, and A. Saifudin, "Pengembangan Kecerdasan Buatan Whatsapp ChatBot untuk Mahasiswa," *J. Teknol. Sist. Inf. dan Apl. ISSN*, vol. 2654, p. 3788, 2022.
- [33] R. A. Susanto and D. Sunandar, "Implementasi Venom Bot Dan Whatsapp Pada Upaya Membumikan Statistik Melalui Asam Pedas," *J. Ilmu Komput. dan Bisnis*, vol. 12, no. 2, pp. 148–158.
- [34] M. Wali *et al.*, *Pengantar 15 Bahasa Pemrograman Terbaik Di Masa Depan (Referensi & Coding Untuk Pemula)*. PT. Sonpedia Publishing Indonesia, 2023.
- [35] S. Rahman, A. Sembiring, D. Siregar, I. G. Prahmana, R. Puspadini, and M. Zen, "Python: Dasar dan Pemrograman Berorientasi Objek," *Penerbit Tahta Media*, 2023.
- [36] M. Putri, D. J. Valentino, and R. Andarsyah, *CARA PRAKTIS MEMBUAT CHATBOT WHATSAPP*. Penerbit Buku Pedia, 2023.
- [37] I. Kurniawan and F. Rozi, "REST API Menggunakan NodeJS pada Aplikasi Transaksi Jasa Elektronik Berbasis Android," *JITSI J. Ilm. Teknol. Sist. Inf.*, vol. 1, no. 4, pp. 127–132, 2020.
- [38] D. F. Ningtyas and N. Setiyawati, "Implementasi Flask Framework pada Pembangunan Aplikasi Purchasing Approval Request," *J. Janitra Inform. dan Sist. Inf.*, vol. 1, no. 1, pp. 19–34, 2021.
- [39] R. K. Ngantung and M. A. I. Pakereng, "Model Pengembangan Sistem Informasi Akademik Berbasis User Centered Design Menerapkan Framework Flask Python," *J. Media Inform. Budidarma*, vol. 5, no. 3, pp. 1052–1062, 2021.
- [40] R. S. Pressman, "Rekayasa perangkat lunak: pendekatan praktisi." Andi, 2012.

- [41] G. W. Sasmito, “Penerapan metode *Waterfall* pada desain sistem informasi geografis industri kabupaten Tegal,” *J. Inform. J. Pengemb. IT*, vol. 2, no. 1, pp. 6–12, 2017.
- [42] “Sistem Informasi Geografis Pemetaan Destinasi Wisata di Kabupaten Sukabumi Berbasis Web”.
- [43] S. Somantri, I. L. Kharisma, and N. Angelina, “Rancang Bangun Aplikasi Pendeteksi Tingkat Kepekatan Asap Hasil Pembakaran Berbasis Internet of Things,” *J-SAKTI (Jurnal Sains Komput. dan Inform.)*, vol. 7, no. 2, pp. 721–730, 2023.
- [44] S. Somantri, G. P. Insany, and R. R. Putra, “Perancangan sistem bimbingan syarat kecakapan umum pramuka berbasis Android,” *IDEALIS Indones. J. Inf. Syst.*, vol. 6, no. 2, pp. 201–210, 2023.
- [45] A. A. Yunanto, W. Yuwono, and J. G. P. Arrochim, “Unit Testing Dan User Review Pada Sistem Informasi Kegiatan Pengabdian Masyarakat Berbasis Website Dan Android,” *POROS Tek.*, vol. 15, no. 1, pp. 30–35, 2023.
- [46] J. Rumbaugh, *The unified modeling language reference manual*. Pearson Education India, 2005.
- [47] G. P. Insany, M. R. A. Rustandi, and A. Fergina, “Sistem Informasi Geografis Pemetaan Fasilitas Kesehatan di Kota Sukabumi Berbasis Web,” *J. RESTIKOM Ris. Tek. Inform. dan Komput.*, vol. 4, no. 1, pp. 1–13, 2022.
- [48] P. T. I. U. P. Y. Nurirwan Saputra, “Pengenalan Data Mining,” 2023.

