

**PERANCANGAN SISTEM *BUSINESS INTELLIGENCE*
BERBASIS *RULE-BASED CHATBOT* DENGAN *TIME SERIES*
FORECASTING DAN TEKNOLOGI BAHASA ALAMI
UNTUK ANALISIS DAN PREDIKSI KINERJA BISNIS**

Proposal Tugas Akhir

Oleh

**Thalita Zahra Sutejo
18222023**



**PROGRAM STUDI SISTEM DAN TEKNOLOGI INFORMASI
SEKOLAH TEKNIK ELEKTRO DAN INFORMATIKA
INSTITUT TEKNOLOGI BANDUNG
Desember 2025**

LEMBAR PENGESAHAN

PERANCANGAN SISTEM *BUSINESS INTELLIGENCE* BERBASIS *RULE-BASED CHATBOT* DENGAN *TIME SERIES* *FORECASTING* DAN TEKNOLOGI BAHASA ALAMI UNTUK ANALISIS DAN PREDIKSI KINERJA BISNIS

Proposal Tugas Akhir

Oleh

Thalita Zahra Sutejo
18222023

Program Studi Sistem dan Teknologi Informasi
Sekolah Teknik Elektro dan Informatika
Institut Teknologi Bandung

Proposal Tugas Akhir ini telah disetujui dan disahkan
di Bandung, pada tanggal 2 Desember 2025

Pembimbing

Dr. Ir. Dimitri Mahayana, M.Eng.

NIP. 196808091991021001

DAFTAR ISI

DAFTAR GAMBAR	v
DAFTAR TABEL	vi
DAFTAR KODE	vii
I PENDAHULUAN	1
I.1 Latar Belakang	1
I.2 Rumusan Masalah	3
I.3 Tujuan	4
I.4 Batasan Masalah	5
I.5 Metodologi	6
I.5.1 Tahapan Investigasi dan Pengumpulan Fakta	6
I.5.2 Studi Literatur	7
I.5.3 Perancangan Sistem	7
I.5.4 Implementasi Sistem	8
I.5.5 Pengujian dan Evaluasi	8
I.5.6 Dokumentasi dan Penyusunan Laporan	9
II STUDI LITERATUR	11
II.1 <i>Business Intelligence</i> dan Arsitektur Sistem	11
II.1.1 Konsep <i>Business Intelligence</i>	11
II.1.2 Arsitektur <i>Business Intelligence</i> Hibrida	12
II.1.3 <i>Semantic Layer</i> dalam <i>Business Intelligence</i>	13
II.1.3.1 Komponen Inti <i>Semantic Layer</i>	13
II.2 <i>Pattern-Based Intent Classification</i> dengan <i>Confidence Level</i>	14
II.2.1 Konsep <i>Intent Classification</i>	14
II.2.2 Klasifikasi <i>Intent</i> Berbasis <i>Transformer</i>	15
II.2.3 <i>Pattern Matching</i> untuk <i>Intent Classification</i>	16
II.2.4 <i>Confidence Score</i> dalam <i>Intent Classification</i>	16
II.2.4.1 <i>Confidence Threshold</i>	17
II.2.5 Evaluasi Performa <i>Intent Classification</i>	17
II.2.5.1 Metrik Evaluasi Klasifikasi <i>Intent</i> Multi-Kelas	17
II.2.5.2 <i>Confusion Matrix</i> dan <i>Confidence Score Analysis</i>	21
II.2.5.3 Target Performa <i>Intent Classification</i>	21
II.3 <i>Template-Based Natural Language Generation</i>	22

II.3.1	Konsep <i>Natural Language Generation</i>	22
II.3.2	<i>Template-Based Natural Language Generation</i>	23
II.3.3	<i>Template Rewriting</i> dengan <i>Pre-trained Language Models</i>	23
II.3.4	Keamanan dan Privasi Data dalam NLG Internal	24
II.3.4.1	<i>Differential Privacy</i> untuk NLG	24
II.3.4.2	Strategi Keamanan Data dalam NLG	24
II.3.4.3	Mitigasi Risiko dalam NLG Internal	25
II.3.5	Evaluasi Kualitas <i>Natural Language Generation</i>	25
II.3.5.1	Metrik Evaluasi NLG	25
II.3.5.2	Metrik <i>Task-Specific</i> untuk BI <i>Chatbot</i>	26
II.3.5.3	Target Performa NLG	26
II.4	<i>Time Series Forecasting</i>	27
II.4.1	Konsep <i>Time Series</i> dan Peramalan	27
II.4.2	Model ARIMA untuk <i>Time Series</i>	27
II.4.3	Model SARIMA untuk Data Musiman	28
II.4.4	Model LSTM untuk <i>Time Series</i>	29
II.4.5	Perbandingan ARIMA, SARIMA, dan LSTM	30
II.4.5.1	Studi Empiris Perbandingan Model	30
II.4.5.2	Model Hibrida ARIMA-LSTM	30
II.4.5.3	Perbandingan Karakteristik Model	31
II.4.6	Evaluasi Performa <i>Time Series Forecasting</i>	32
II.4.6.1	Metrik Evaluasi <i>Forecasting</i>	32
II.4.6.2	<i>Residual Analysis</i>	33
II.4.6.3	<i>Cross-Validation</i> untuk <i>Time Series</i>	33
II.4.6.4	Target Performa <i>Forecasting</i>	33
II.5	Integrasi Sistem dan Penelitian Terkait	34
II.5.1	Arsitektur <i>Chatbot</i> Hibrida	34
II.5.1.1	Komponen Arsitektur <i>Chatbot</i> Modern	34
II.5.1.2	<i>Chatbot</i> Hibrida dengan AI dan <i>Human Oversight</i>	35
II.5.1.3	Pasar <i>Chatbot</i> AI Global	35
II.5.2	Integrasi <i>Pattern-Based Intent</i> dan <i>Template-Based</i> NLG	36
II.5.2.1	Alur Kerja Terintegrasi	36
II.5.2.2	Evaluasi Sistem Terintegrasi	37
II.5.3	<i>Customer Churn Prediction</i> dengan <i>Machine Learning</i>	37
II.5.3.1	Model <i>Machine Learning</i> untuk <i>Churn Prediction</i>	37
II.5.3.2	Evaluasi Model <i>Churn Prediction</i>	38
II.5.3.3	Fitur-fitur untuk <i>Churn Prediction</i>	38

II.5.3.4	<i>Real-Time Churn Prediction</i>	38
II.5.4	Integrasi <i>Churn Prediction</i> dengan <i>Chatbot</i> BI	39
III	ANALISIS MASALAH	40
III.1	Analisis Kondisi Saat Ini	40
III.2	Analisis Kebutuhan	43
III.2.1	Identifikasi Masalah Pengguna	43
III.2.2	Kebutuhan Fungsional	45
III.2.3	Kebutuhan Nonfungsional	49
III.3	Analisis Pemilihan Solusi	53
III.3.1	Alternatif Solusi	54
III.3.2	Analisis Penentuan Solusi	59

DAFTAR GAMBAR

I.1	Metodologi <i>CRISP-DM</i>	6
II.1	<i>Pattern-Based Intent Classification</i>	14
II.2	<i>Natural Language Generation</i>	22
II.3	<i>Time Series Forecasting</i>	27

DAFTAR TABEL

II.1	Perbandingan Metode Peramalan Deret Waktu	31
III.1	Kebutuhan Fungsional Sistem <i>Chatbot</i> BI	46
III.2	Kebutuhan Nonfungsional Sistem <i>Chatbot</i> BI	49
III.3	Perbandingan Alternatif Solusi Teknologi	54
III.4	Analisis <i>Balanced Scorecard</i> untuk Aturan Sederhana	60
III.5	Analisis <i>Balanced Scorecard</i> untuk BI Komprehensif	62
III.6	Analisis <i>Balanced Scorecard</i> untuk BI Tradisional	66
III.7	Analisis <i>Balanced Scorecard</i> untuk LLM	68
III.8	Perbandingan <i>Balanced Scorecard</i> Alternatif Solusi	72

DAFTAR KODE

BAB I

PENDAHULUAN

I.1 Latar Belakang

Transformasi digital di sektor bisnis telah mengubah cara perusahaan menganalisis dan memanfaatkan data untuk pengambilan keputusan strategis. *Business Intelligence* (BI) kini menjadi tulang punggung operasional organisasi modern yang ingin tetap kompetitif di pasar global. Pasar global *Business Intelligence* diproyeksikan mencapai USD 36,82 miliar pada tahun 2025 dan tumbuh menjadi USD 116,25 miliar pada tahun 2033, dengan tingkat pertumbuhan tahunan gabungan (CAGR) sebesar 14,98% (Straits Research 2024). Pertumbuhan ini menunjukkan bahwa organisasi di seluruh dunia semakin menyadari nilai strategis sistem analitik berbasis data untuk meningkatkan efisiensi operasional dan daya saing.

Di Indonesia, adopsi teknologi kecerdasan buatan (*Artificial Intelligence/AI*) dan digitalisasi juga mengalami akselerasi signifikan. Pemerintah Indonesia memproyeksikan bahwa AI akan menghasilkan manfaat ekonomi hingga USD 366 miliar selama dekade mendatang (Intimedia Desember 2024). Selain itu, menurut laporan *East Ventures—Digital Competitiveness Index 2025*, lebih dari 80% pelaku bisnis di Indonesia telah mengintegrasikan AI ke dalam operasi bisnis, meskipun hanya sekitar 13% yang menggunakan AI pada tingkat lanjutan (East Ventures July 2025).

Salah satu tantangan utama dalam implementasi BI adalah kompleksitas akses dan interpretasi data bagi pengguna nonteknis. Sistem BI konvensional sering memerlukan keahlian khusus dalam bahasa kueri (seperti SQL) dan pemahaman mendalam tentang struktur basis data sehingga menjadi hambatan bagi manajer dan staf operasional yang membutuhkan *insight* cepat untuk keputusan harian. Untuk mengatasi hal tersebut, teknologi *Conversational AI* berbasis *chatbot* muncul sebagai solusi yang menjanjikan. Pasar *Conversational AI* global diperkirakan mencapai USD 49,7 miliar pada tahun 2025 dan tumbuh dengan CAGR sekitar 23,6% hingga 2032

(Qaltivate 2025).

Implementasi *chatbot* dalam konteks layanan pelanggan dan operasional internal telah menunjukkan hasil yang positif. Ringkasan industri menunjukkan bahwa *chatbot* dapat menjawab hingga 80% pertanyaan standar pelanggan, mempercepat waktu respons, dan memangkas beban kerja agen manusia (Fullview 2025). Dalam konteks BI, *chatbot* berfungsi sebagai antarmuka percakapan yang memungkinkan pengguna internal, seperti manajer penjualan, tim layanan pelanggan, dan analis bisnis, mengakses data analitik, laporan, dan prediksi melalui kueri bahasa alami tanpa keahlian teknis. Studi (Iqbal, Ahmad, dan Khan 2021) menunjukkan bahwa penerapan *natural language processing* (NLP) pada *chatbot* meningkatkan efisiensi dan akurasi pelayanan pelanggan secara signifikan.

Teknologi NLP menjadi fondasi yang memungkinkan sistem memahami dan merespons kueri pengguna dalam bahasa alami. Pasar NLP global diproyeksikan tumbuh dari USD 35,43 miliar pada tahun 2024 menjadi USD 438,08 miliar pada tahun 2034, dengan CAGR sekitar 28,6% (Precedence Research 2025). Dalam aplikasi BI, NLP memungkinkan sistem menginterpretasikan *intent* pengguna, mengekstraksi entitas penting dari kueri, dan menghasilkan respons relevan berdasarkan konteks percakapan, baik melalui pendekatan *rule-based* untuk kueri terstruktur maupun melalui model pembelajaran mesin untuk kueri yang lebih kompleks.

Selain kemampuan analisis deskriptif dan diagnostik, kebutuhan akan kemampuan prediktif dalam BI juga meningkat. Analisis prediktif memungkinkan organisasi mengantisipasi permintaan layanan pelanggan, merencanakan alokasi sumber daya, dan mengidentifikasi potensi *churn* sebelum terjadi. Dalam konteks ini, permintaan layanan pelanggan didefinisikan sebagai minat, pesanan, dan kebutuhan pelanggan terhadap produk atau layanan yang ditawarkan perusahaan, yang dapat diukur melalui data pemesanan, target penjualan, tingkat *churn*, dan pendapatan pelanggan. Pasar *time series forecasting* secara global diperkirakan tumbuh dari sekitar USD 11,17 miliar (2024) menjadi USD 36,9 miliar pada 2032, dengan CAGR ~16,12% (WiseGuy Reports 2024).

Metode *forecasting* yang lazim meliputi model statistik dan model *deep learning* untuk menangkap pola temporal. Pada kasus prediksi *customer churn* di industri telekomunikasi, baik Random Forest maupun LSTM efektif dalam prediksi *churn*, tetapi kegunaannya bergantung pada sifat data dan batasan bisnis. LSTM lebih baik bila interaksi pelanggan historis dan tren kepuasan menjadi kunci dalam pengambilan keputusan (Mena dkk. 2023).

Integrasi tiga komponen utama, *Rule-Based Query*, *Natural Language Generation*, dan *Time Series Forecasting*, ke dalam satu sistem BI berbasis *chatbot* menawarkan solusi komprehensif bagi kebutuhan analisis dan prediksi di lingkungan bisnis modern. Sistem demikian berpotensi memperkuat pengambilan keputusan proaktif berbasis data serta meningkatkan efisiensi operasional internal perusahaan.

I.2 Rumusan Masalah

Sebagian besar sistem *Business Intelligence* untuk analisis indikator permintaan terkait pelanggan di lingkungan internal perusahaan pada saat ini masih mengandalkan pelaporan manual, belum terintegrasi dengan antarmuka percakapan berbasis *chatbot*, dan minim pemanfaatan kecerdasan buatan untuk analisis prediktif serta interaksi berbasis bahasa alami. Kondisi ini memunculkan sejumlah permasalahan sebagai berikut.

1. Keterbatasan akses analitik berbasis percakapan

Tidak tersedia sistem yang mampu memproses dan memahami permintaan data serta analisis bisnis secara otomatis dari pengguna internal melalui percakapan dalam bahasa alami. Hal ini menyebabkan pegawai mengalami kendala dalam memperoleh *insight* terkait pesanan (*order*), pencapaian target penjualan, *churn* pelanggan, serta pendapatan secara cepat dan intuitif untuk mendukung pengambilan keputusan harian.

2. Ketiadaan fitur prediksi yang mudah diakses pengguna nonteknis

Belum terdapat fitur prediksi indikator utama bisnis terkait pelanggan, seperti peramalan jumlah pesanan, *churn* pelanggan, dan target pendapatan berbasis pembelajaran mesin atau peramalan deret waktu yang dapat diakses oleh pengguna nonteknis. Keadaan ini menyebabkan perencanaan kapasitas, evaluasi target, dan penetapan strategi bisnis perusahaan menjadi kurang responsif terhadap tren serta dinamika bisnis pada masa yang akan datang.

3. Minimnya integrasi *rule-based query* dan *natural language generation*

Integrasi *rule-based query* dan pemrosesan bahasa alami (*natural language generation*) dalam sistem BI internal masih sangat terbatas sehingga proses pencarian data, penyaringan metrik utama, serta penyusunan laporan kinerja bisnis masih harus dilakukan secara manual dan berulang. Keadaan tersebut menimbulkan keterlambatan pengambilan keputusan serta penyajian data yang kurang selaras dengan kebutuhan analisis oleh manajemen.

Bagaimana merancang dan mengembangkan sistem *Business Intelligence* berbasis *chatbot* yang terintegrasi dengan teknologi *rule-based query*, pemrosesan bahasa

alami dengan (*natural language generation*), dan algoritma peramalan deret waktu sehingga seluruh pengguna internal perusahaan dapat menganalisis, memantau, dan memprediksi indikator bisnis pelanggan, seperti pesanan, target, *churn*, dan pendapatan melalui antarmuka percakapan, serta memperoleh *insight* berbasis data yang akurat, efisien, dan relevan untuk mendukung pengambilan keputusan bisnis dan perencanaan strategis perusahaan?

I.3 Tujuan

Tujuan utama tugas akhir ini adalah merancang dan mengembangkan sistem *Business Intelligence* berbasis *chatbot* yang terintegrasi dengan teknologi *rule-based query*, pemrosesan bahasa alami (*Natural Language Generation/NLG*), dan algoritma peramalan deret waktu (*time series forecasting*) untuk memfasilitasi pengguna internal perusahaan dalam menganalisis dan memprediksi indikator bisnis pelanggan.

Secara terperinci, tujuan yang hendak dicapai dalam tugas akhir ini adalah sebagai berikut.

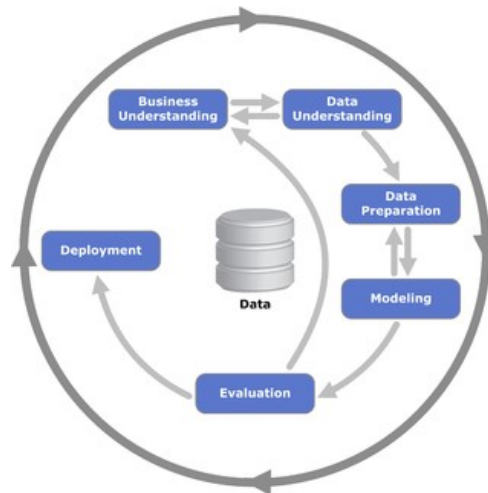
1. Merancang arsitektur sistem *Business Intelligence* berbasis *chatbot* yang mampu memproses permintaan analisis data dan memberikan *insight* berbasis data kepada pengguna internal perusahaan secara otomatis.
2. Mengimplementasikan modul pemrosesan bahasa alami sehingga pengguna internal dapat mengajukan pertanyaan dan permintaan analisis bisnis menggunakan bahasa alami melalui antarmuka *chatbot*.
3. Mengimplementasikan modul mekanisme *rule-based query* untuk menangani permintaan data yang bersifat terstruktur dan berulang dengan respons yang cepat dan konsisten.
4. Mengintegrasikan model peramalan deret waktu (*time series forecasting*) untuk menghasilkan prediksi terhadap indikator bisnis pelanggan, seperti volume pesanan, pencapaian target penjualan, tingkat *churn* pelanggan, dan proyeksi pendapatan.
5. Mengintegrasikan model yang menghasilkan bahasa tertulis atau lisan alami dari data terstruktur.
6. Merancang antarmuka percakapan yang dapat menampilkan percakapan.
7. Menguji dan mengevaluasi kinerja sistem dari segi akurasi prediksi, kecepatan respons, serta kemudahan penggunaan (*usability*) oleh pengguna internal perusahaan.

I.4 Batasan Masalah

Dalam memfokuskan ruang lingkup penelitian dan memastikan kelayakan pelaksanaan tugas akhir, batasan-batasan penelitian ditetapkan sebagai berikut.

1. Sistem dirancang khusus untuk pengguna internal perusahaan (manajer, analis bisnis, dan staf operasional), bukan untuk pelanggan eksternal.
2. Sistem *chatbot* hanya mendukung interaksi dalam Bahasa Indonesia dan Bahasa Inggris.
3. Sistem menggunakan data yang terbatas pada data historis pelanggan yang mencakup informasi pesanan (*order*), target penjualan, tingkat *churn*, dan pendapatan dalam kurun waktu minimal dua tahun terakhir.
4. Sistem hanya menangani permintaan terkait analisis deskriptif, diagnostik, dan prediktif terhadap indikator bisnis pelanggan. Sistem tidak mencakup fungsi transaksional atau modifikasi data.
5. Sistem menggunakan metode peramalan deret waktu berbasis deep learning yang dibatasi pada algoritma LSTM, GRU, Temporal Fusion Transformer (TFT), atau kombinasi dari algoritma-algoritma tersebut.
6. Sistem dirancang untuk mengenali maksimal 20 hingga 30 *intent* utama yang berkaitan dengan analisis dan prediksi indikator bisnis pelanggan.
7. Sistem yang dikembangkan merupakan purwarupa mandiri (*standalone*) dan belum terintegrasi secara penuh dengan sistem ERP (*Enterprise Resource Planning*) atau CRM (*Customer Relationship Management*) yang ada di perusahaan.
8. Sistem diimplementasikan pada platform berbasis *web* dengan menggunakan *framework* yang umum untuk pengembangan *chatbot*.
9. Sistem dievaluasi melalui pengujian fungsional, pengukuran akurasi prediksi, analisis waktu respons, serta survei kepuasan pengguna dengan jumlah responden terbatas (minimal 20 hingga 30 pengguna internal).
10. Sistem mengimplementasikan fitur keamanan dan privasi data dibatasi pada tingkat dasar, seperti autentikasi pengguna dan enkripsi koneksi, tanpa mencakup standar keamanan tingkat *enterprise* seperti ISO 27001.

I.5 Metodologi



Gambar I.1 Metodologi *CRISP-DM*

Metodologi pelaksanaan tugas akhir ini mengadopsi kerangka kerja hibrida yang menggabungkan CRISP-DM (*Cross-Industry Standard Process for Data Mining*) untuk tahapan pengolahan data dan analitik. Pendekatan ini dipilih agar proses penelitian berjalan secara iteratif, luwes, dan tetap sistematis dalam menyelesaikan permasalahan yang telah dirumuskan.

CRISP-DM (*Cross-Industry Standard Process for Data Mining*) adalah metodologi standar lintas industri untuk penambangan data yang menyediakan pendekatan terstruktur dalam enam fase utama: pemahaman bisnis (*business understanding*), pemahaman data (*data understanding*), persiapan data (*data preparation*), pemodelan (*modeling*), evaluasi (*evaluation*), dan penyebaran (*deployment*) (DataScience-PM, 2024). Metodologi ini bersifat iteratif, yang memungkinkan pengulangan fase jika diperlukan untuk meningkatkan kualitas hasil. Dalam penelitian ini, **CRISP-DM digunakan** untuk mengelola aspek penambangan data dan analitik prediktif (khususnya pengembangan model peramalan deret waktu dan NLP).

I.5.1 Tahapan Investigasi dan Pengumpulan Fakta

Tahapan ini bertujuan untuk mengumpulkan data dan informasi faktual sebagai landasan perumusan masalah. Kegiatan yang dilakukan meliputi:

- a. **Observasi sistem yang ada**, yaitu pengamatan terhadap sistem *Business Intelligence* yang saat ini digunakan organisasi untuk mengidentifikasi keterbatasan dan kebutuhan pengguna internal.

- b. **Wawancara dengan pemangku kepentingan**, dilakukan secara terstruktur dengan minimal lima hingga sepuluh pemangku kepentingan internal (manajer layanan pelanggan, analis data, staf operasional) untuk menggali kebutuhan fungsional dan ekspektasi terhadap sistem baru.
- c. **Survei kebutuhan pengguna**, disebarkan kepada minimal tiga puluh hingga lima puluh calon pengguna untuk mengidentifikasi kebutuhan informasi, preferensi interaksi, dan kebutuhan analitik.
- d. **Analisis data historis**, dilakukan untuk menelaah volume dan tren permintaan, pola musiman, serta kelengkapan dan kualitas data yang tersedia.
- e. **Dokumentasi temuan**, mengompilasi seluruh hasil observasi, wawancara, survei, dan analisis dalam bentuk laporan yang akan dilampirkan.

I.5.2 Studi Literatur

Tahapan ini bertujuan untuk mengumpulkan, mengelompokkan, dan menyaring literatur yang relevan dengan topik penelitian. Proses yang dilakukan meliputi:

- a. **Pencarian literatur sistematis** pada basis data akademik (IEEE Xplore, ScienceDirect, SpringerLink, ACM Digital Library, Google Scholar) menggunakan kata kunci terkait *business intelligence*, *chatbot*, *natural language generation*, dan *time series forecasting*. Kriteria inklusi adalah artikel telaah sejawat (*peer-reviewed*) tahun 2020–2025 yang relevan dengan topik.
- b. **Pengelompokan literatur** berdasarkan tema utama: konsep *Business Intelligence*, teknologi *chatbot* dan *conversational AI*, pemrosesan bahasa alami, metode peramalan deret waktu, integrasi sistem, dan studi kasus implementasi.
- c. **Penapisan dan analisis kualitas** literatur berdasarkan faktor dampak, jumlah sitasi, metodologi, dan relevansi terhadap penelitian.
- d. **Sintesis literatur** untuk membangun kerangka teoretis, mengidentifikasi praktik terbaik dan teknologi terkini, serta menyusun tinjauan pustaka pada Bab II.
- e. **Dokumentasi proses** dalam bentuk matriks literatur dan pengelolaan referensi menggunakan *reference manager*.

I.5.3 Perancangan Sistem

Tahapan ini meliputi perancangan arsitektur dan komponen sistem secara terperinci, mencakup:

- a. **Perancangan arsitektur sistem**, yaitu penyusunan arsitektur berlapis yang

mengintegrasikan *chatbot*, pemrosesan bahasa alami, mesin *rule-based query*, dan peramalan deret waktu, serta penentuan teknologi dan kerangka kerja yang akan digunakan.

- b. **Perancangan modul *chatbot***, meliputi komponen pemahaman bahasa alami (*Natural Language Understanding*), manajemen dialog (*Dialogue Management*), pembangkitan bahasa alami (*Natural Language Generation*), serta pendefinisian *intent* dan entitas yang akan dikenali.
- c. **Pengimplementasian *rule-based query***, yaitu pendefinisian aturan untuk kueri terstruktur dan berulang serta pemetaan antara *intent* dengan templat kueri.
- d. **Perancangan model peramalan deret waktu**, meliputi penentuan variabel target, perancangan praolah data, pemilihan arsitektur model (ARIMA, SARIMA, LSTM), dan strategi evaluasi model.
- e. **Perancangan basis data dan gudang data**, yaitu perancangan skema basis data untuk menyimpan data pelanggan, hasil prediksi, dan data pelatihan NLP, serta perancangan proses ETL (*Extract, Transform, Load*).
- f. **Perancangan antarmuka pengguna**, meliputi pembuatan *wireframe* dan *mockup* untuk antarmuka *chatbot*.

I.5.4 Implementasi Sistem

Tahapan ini meliputi pengembangan sistem berdasarkan rancangan yang telah ditetapkan, mencakup:

- a. **Pengembangan *backend***, yaitu implementasi mesin *rule-based query*, model NLP untuk klasifikasi *intent* dan pengenalan entitas, model peramalan deret waktu, serta penyediaan API *backend*.
- b. **Pengembangan *frontend***, meliputi implementasi antarmuka *chatbot*.
- c. **Pengembangan basis data**, yaitu implementasi skema basis data, proses ETL untuk integrasi data, dan *data pipeline* untuk pemrosesan data.
- d. **Integrasi komponen**, meliputi integrasi *chatbot* dengan mesin *rule-based query* dan NLP, integrasi model peramalan dengan sistem BI, serta implementasi mekanisme tembolok (*caching*) untuk performa.

I.5.5 Pengujian dan Evaluasi

Tahapan ini bertujuan menilai kinerja sistem secara menyeluruh—mencakup aspek fungsionalitas, akurasi, performa, dan kebergunaan.

- a. **Pengujian unit**, yaitu setiap komponen diuji secara terpisah agar fungsinya terverifikasi dengan benar dan stabil.

- b. **Pengujian integrasi**, yaitu Interaksi antarmodul diverifikasi melalui skenario ujung-ke-ujung, dari masukan percakapan hingga keluaran berupa kueri yang dieksekusi, prediksi, narasi, dan visualisasi.
- c. **Pengujian fungsional**, yaitu Pemenuhan seluruh kebutuhan fungsional divalidasi menggunakan himpunan skenario representatif yang mencerminkan pertanyaan bisnis prioritas.
- d. **Evaluasi model NLP (klasifikasi intent multikelas, satu label per ujaran)**, yaitu hasil dilaporkan sebagai *top-1* dan *top-k* accuracy; *precision*, *recall*, dan *F1-score* per kelas yang kemudian dirata-ratakan dengan pendekatan *macro*, *micro*, dan *weighted*. Disajikan pula *confusion matrix* berukuran $K \times K$, evaluasi ambang kepercayaan untuk mekanisme *fallback*/UNKNOWN, serta ringkasan kalibrasi probabilitas.
- e. **Evaluasi model peramalan deret waktu**, yaitu protokol yang digunakan ialah *rolling-origin backtesting* pada beberapa horizon prediksi dengan *seasonal naive* sebagai garis dasar. Laporan mencakup MAE, RMSE, sMAPE, dan MASE, dilengkapi *skill score* terhadap garis dasar. Hasil disajikan per horizon dan per entitas prioritas (median dan rentang antar kuartil).
- f. **Evaluasi performa sistem**, yaitu waktu respons, *throughput*, dan latensi diukur pada beban realistis; *load testing* dilakukan untuk menilai skalabilitas.
- g. **Evaluasi kebergunaan (usability)**, yaitu *User Acceptance Testing* melibatkan sedikitnya 20–30 pengguna internal. Hasil dirangkum menggunakan skor *System Usability Scale* (SUS) dan umpan balik kualitatif terkait kemudahan, kejelasan, dan alur kerja.

I.5.6 Dokumentasi dan Penyusunan Laporan

Tahapan akhir meliputi dokumentasi dan penyusunan laporan tugas akhir, mencakup:

- a. **Dokumentasi teknis**, yaitu penyusunan dokumentasi API, dokumentasi kode, panduan pengguna (*user manual*), dan panduan penyebaran (*deployment guide*).
- b. **Penyusunan laporan tugas akhir**, meliputi penulisan laporan sesuai format yang ditetapkan dengan penyuntingan dan pemeriksaan kelengkapan.
- c. **Persiapan presentasi**, meliputi pembuatan *slide* presentasi, penyusunan skenario demonstrasi sistem, dan kurasi materi pendukung.
- d. **Revisi dan finalisasi**, yaitu revisi berdasarkan masukan pembimbing, finalisasi laporan dan sistem, serta persiapan sidang tugas akhir.

Seluruh proses metodologi ini akan didokumentasikan dengan baik untuk memastikan keterbukaan, kemampuan direproduksi, dan akuntabilitas penelitian.

BAB II

STUDI LITERATUR

Bab ini menyajikan kajian teoretis dan tinjauan pustaka yang relevan dengan topik penelitian. Pembahasan meliputi konsep *Business Intelligence*, teknologi *chatbot* dan *Conversational AI*, klasifikasi *intent* berbasis pola, pembangkitan bahasa alami berbasis templat, serta peramalan deret waktu yang mendukung analisis dan prediksi permintaan layanan pelanggan. Kajian literatur ini disusun secara sistematis untuk memberikan landasan teoretis yang kuat bagi pengembangan sistem yang diusulkan, dengan fokus pada integrasi *pattern-based intent classification*, *template-based natural language generation*, dan *time series forecasting* dalam satu platform terintegrasi.

II.1 *Business Intelligence* dan Arsitektur Sistem

Bagian ini membahas konsep dasar *Business Intelligence*, evolusinya, serta arsitektur sistem yang mendukung implementasi BI modern. Pembahasan mencakup komponen-komponen utama arsitektur BI dan peran *semantic layer* dalam menjembatani kesenjangan antara data teknis dan kebutuhan bisnis. Pemahaman yang komprehensif tentang arsitektur BI menjadi fondasi penting dalam merancang sistem yang mampu mengintegrasikan berbagai komponen teknologi secara efektif.

II.1.1 Konsep *Business Intelligence*

Business Intelligence (BI) merupakan sistem teknologi yang mengumpulkan, mengorganisasi, dan menganalisis data dari berbagai sumber dalam suatu bisnis untuk membantu perusahaan mengubah data mentah menjadi wawasan yang berguna sehingga mereka dapat membuat keputusan yang lebih baik (Sorour 2020). Evolusi BI dari sistem berbasis aturan tradisional ke platform berbasis kecerdasan buatan telah meningkatkan secara signifikan analitik prediktif, otomatisasi, dan kemampuan pengambilan keputusan waktu nyata di seluruh industri (Khaddam dan Alzghoul

2025).

Implementasi BI di institusi pendidikan tinggi telah menunjukkan kemampuan untuk memantau aktivitas jaminan kualitas dan mendukung pengambilan keputusan strategis (Sorour 2020). Penelitian Shah (2025) menunjukkan bahwa organisasi yang mengimplementasikan arsitektur analitik hibrida melaporkan peningkatan kemampuan penanganan data sebesar 85% dibandingkan sistem tradisional, dengan konsistensi dan standar kualitas data yang terjaga. Pasar BI global terus menunjukkan pertumbuhan yang signifikan, dengan proyeksi mencapai USD 53,8 miliar pada tahun 2033, meningkat dari USD 25,9 miliar pada tahun 2024, dengan CAGR sebesar 8,48% (Straits Research 2024). Pertumbuhan ini menunjukkan relevansi dan kebutuhan akan solusi BI yang inovatif dalam menghadapi kompleksitas data bisnis modern.

II.1.2 Arsitektur *Business Intelligence* Hibrida

Arsitektur BI hibrida mengintegrasikan sistem BI tradisional dengan kemampuan kecerdasan buatan untuk meningkatkan pengambilan keputusan, analitik prediktif, dan efisiensi operasional. Pendekatan ini menyajikan struktur yang memanfaatkan model pembelajaran mesin bersama pelaporan BI tradisional untuk menjembatani kesenjangan antara analisis historis dan wawasan berbasis data waktu nyata (Shah 2025).

Komponen utama arsitektur BI hibrida mencakup beberapa lapisan yang saling terintegrasi (Shah 2025). Lapisan sumber data mencakup sumber internal seperti sistem CRM dan perangkat lunak keuangan, serta sumber eksternal seperti data media sosial dan laporan pasar. Gudang data berfungsi sebagai fasilitas penyimpanan terpusat untuk data yang terorganisasi dan siap digunakan untuk pelaporan. Mesin analitik bertanggung jawab untuk pengenalan pola dan analisis tren, memberikan wawasan prediktif untuk perencanaan strategis. Lapisan visualisasi menyediakan dashboard interaktif dan laporan yang meningkatkan adopsi pengguna melalui presentasi data yang intuitif.

Implementasi arsitektur analitik hibrida menghasilkan peningkatan 27% dalam proses pengambilan keputusan dan peningkatan 31% dalam metrik kinerja organisasi secara keseluruhan (Shah 2025). Integrasi komponen-komponen ini membentuk fondasi yang kuat untuk sistem BI modern yang responsif dan dapat diandalkan. Keberhasilan implementasi arsitektur hibrida ini membuka jalan bagi pengembangan sistem yang lebih kompleks dan terintegrasi dengan teknologi percakapan.

II.1.3 *Semantic Layer* dalam *Business Intelligence*

Semantic layer adalah antarmuka berorientasi bisnis yang menjembatani kesenjangan antara model data yang kompleks dan pengguna bisnis, bertindak sebagai lapisan abstraksi yang menerjemahkan struktur data teknis ke dalam istilah dan konsep bisnis yang familiar (Databricks 2025). Lapisan semantik menyediakan pandangan bisnis terpadu terhadap data di seluruh organisasi, terlepas dari lokasi data atau bagaimana secara teknis terstruktur.

Lapisan semantik menyederhanakan konsep dan teknis data untuk pengguna bisnis sehingga mereka tidak perlu mengubah data bisnis yang mendasarinya untuk bekerja dengan cara baru (AtScale 2025). Lapisan ini memungkinkan pengguna bisnis untuk berinteraksi dengan data menggunakan terminologi yang mereka pahami tanpa perlu memahami struktur teknis basis data yang mendasarinya. Kemampuan ini sangat penting dalam konteks sistem *chatbot* BI, dengan memfasilitasi interpretasi kueri bahasa alami menjadi operasi basis data yang sesuai.

II.1.3.1 *Komponen Inti Semantic Layer*

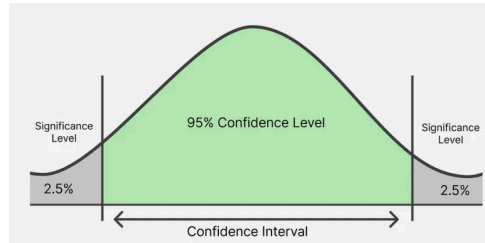
Lapisan semantik memiliki lima komponen inti yang bertindak sebagai fondasi struktural dan teknis (Labs 2024). Komponen pertama adalah definisi model semantik yang menciptakan representasi logis dari domain bisnis, memetakan struktur basis data teknis ke konsep bisnis. Alih-alih bekerja dengan tabel mentah seperti `usr_tbl` atau `trx_hist`, entitas seperti *Customer* atau *Order* didefinisikan untuk merangkum kompleksitas yang mendasarinya.

Komponen kedua adalah manajemen metadata yang menangani informasi tentang data, seperti deskripsi bidang, garis keturunan data (*data lineage*), frekuensi pembaruan, dan metrik kualitas. Lapisan logika bisnis sebagai komponen ketiga berisi aturan perhitungan spesifik untuk metrik bisnis. Komponen keempat adalah lapisan akses data yang menerjemahkan permintaan bisnis menjadi kueri basis data yang dioptimalkan, menerapkan filter keamanan yang diperlukan. Komponen terakhir adalah mekanisme *caching* yang memeriksa apakah hasil perhitungan sudah tersedia dalam tembolok sebelum menjalankan kueri.

Implementasi *semantic layer* yang efektif memungkinkan organisasi untuk mencapai konsistensi dalam pelaporan dan analisis, sekaligus mengurangi kompleksitas teknis yang dihadapi oleh pengguna bisnis. Pemahaman tentang *semantic layer* ini menjadi dasar penting dalam merancang sistem klasifikasi *intent* yang akan dibahas

pada bagian selanjutnya.

II.2 *Pattern-Based Intent Classification dengan Confidence Level*



Gambar II.1 *Pattern-Based Intent Classification*

Bagian ini membahas pendekatan klasifikasi *intent* berbasis pola untuk sistem *chatbot*, dengan fokus pada mekanisme pencocokan pola, penggunaan *confidence score*, dan evaluasi performa. Klasifikasi *intent* merupakan komponen kritis dalam memahami maksud pengguna dari kueri bahasa alami dan menentukan respons yang sesuai berdasarkan data internal perusahaan.

II.2.1 Konsep *Intent Classification*

Klasifikasi *intent* adalah proses menentukan niat atau tujuan pengguna dari masukan mereka, yang merupakan tugas fundamental dalam sistem *chatbot* percakapan (Ouaddi, Benabbou, dan Sael 2025). Inti dari *chatbot* berbasis AI adalah komponen NLU yang mengklasifikasikan *intent* pengguna untuk menghasilkan respons yang sesuai. Tugas klasifikasi ini sangat penting karena menentukan alur percakapan dan tindakan yang akan diambil sistem.

Pasar *Conversational AI* global diproyeksikan mencapai USD 49,9 miliar pada tahun 2030, meningkat dari USD 10,7 miliar pada tahun 2023, dengan CAGR sebesar 25,2% (Qaltivate 2025). Pertumbuhan eksponensial ini didorong oleh meningkatnya adopsi teknologi AI dalam layanan pelanggan dan operasional bisnis. *Chatbot* dapat menjawab hampir 80% dari semua pertanyaan standar, yang secara signifikan mengurangi beban kerja agen manusia dan mempercepat waktu respons (Fullview 2025). Statistik ini menunjukkan potensi besar klasifikasi *intent* dalam meningkatkan efisiensi operasional dan kepuasan pengguna.

II.2.2 Klasifikasi *Intent* Berbasis *Transformer*

Klasifikasi *intent* berbasis *Transformer* merupakan pendekatan modern yang memanfaatkan arsitektur jaringan saraf dengan mekanisme perhatian diri (*self-attention*) untuk memahami konteks ujaran pengguna secara dua arah dan menyeluruh (Merugu dkk. 2024). Model-model seperti BERT (*Bidirectional Encoder Representations from Transformers*), RoBERTa, IndoBERT, dan berbagai *sentence transformer* telah menunjukkan kinerja yang unggul pada tugas klasifikasi teks, termasuk klasifikasi *intent* dalam sistem percakapan (Ozerova 2022; Gehweiler 2024). Pendekatan ini merepresentasikan setiap kalimat sebagai vektor berdimensi tinggi yang memuat informasi semantik kaya, kemudian menambahkan lapisan klasifikasi di bagian akhir untuk memetakan representasi tersebut ke label *intent* yang telah didefinisikan.

Berbagai penelitian melaporkan bahwa model berbasis *Transformer* mampu melampaui pendekatan tradisional seperti LSTM, CNN, maupun metode berbasis fitur manual dalam hal akurasi, skor F1, dan kemampuan menangkap nuansa konteks (Gehweiler 2024). Studi yang membandingkan BERT, RoBERTa, dan IndoBERT untuk klasifikasi *intent chatbot* berbahasa Indonesia menunjukkan bahwa IndoBERT dapat mencapai akurasi hingga sekitar 0,94 berkat proses praprelatihan yang selaras dengan struktur bahasa Indonesia dan korpus yang digunakan. Penelitian lain yang memanfaatkan MiniLM dan RoBERTa untuk modul deteksi *intent* dalam asisten percakapan melaporkan peningkatan akurasi 10–15 persen dibandingkan solusi sebelumnya, sekaligus menekankan pentingnya memperhatikan kecepatan inferensi dan kebutuhan perangkat keras saat model *Transformer* diintegrasikan ke dalam sistem produksi (Ozerova 2022).

Perkembangan model bahasa besar berbasis *Transformer* juga mendorong kajian terbaru yang mengevaluasi penggunaan *sentence transformer*, model BERT terkontrasif, dan bahkan model bahasa besar generatif untuk tugas deteksi *intent* pada agen percakapan tugas-orientasi (Merugu dkk. 2024; Gehweiler dkk. 2025). Hasil kajian tersebut menyoroti adanya kompromi yang jelas antara kualitas prediksi, kemampuan menangani *intent* di luar cakupan (*out-of-scope*), latensi sistem, serta konsumsi sumber daya komputasi, sehingga pemilihan arsitektur perlu mempertimbangkan kebutuhan dan batasan operasional sistem yang dibangun (Merugu dkk. 2024).

Dalam konteks penelitian ini, klasifikasi *intent* berbasis *Transformer* diposisikan sebagai pendekatan yang relevan dan dikaji secara teoretis, tetapi tidak diimplementasikan secara langsung pada prototipe sistem. Pertimbangan utama adalah ruang lingkup pertanyaan yang relatif terbatas pada kebutuhan internal, pola kueri yang

cenderung berulang, serta prioritas untuk menjaga efisiensi pemrosesan dan menghemat penggunaan sumber daya komputasi berbiaya tinggi seperti unit pemroses grafis (GPU) (Ozerova 2022; Gehweiler 2024). Oleh karena itu, penelitian ini memilih pendekatan klasifikasi *intent* berbasis pola dengan mekanisme nilai keyakinan (*pattern-based intent classification with confidence scoring*) yang lebih ringan secara komputasi, sambil menjadikan pendekatan berbasis *Transformer* sebagai acuan pengembangan lanjutan ketika cakupan *intent* dan beban sistem meningkat di masa mendatang.

II.2.3 *Pattern Matching untuk Intent Classification*

Pendekatan berbasis pola (*pattern matching*) dalam klasifikasi *intent* menggunakan aturan dan templat yang telah ditentukan sebelumnya untuk mencocokkan masukan pengguna dengan kategori *intent* yang sesuai. Sistem ini menyimpan pola-pola pertanyaan atau *utterances* untuk setiap *intent*, kemudian membandingkan input pengguna dengan pola-pola yang telah didefinisikan untuk menemukan kecocokan terbaik.

Pendekatan *pattern matching* efektif untuk menangani kueri terstruktur dan berulang dalam domain spesifik. Implementasi dapat dilakukan menggunakan ekspresi reguler, pencocokan *n-gram*, atau algoritma kesamaan string seperti *Fuzzy String Matching* (Mulyatun, Kurniawan, dan Hudaya 2021). Keuntungan utama pendekatan ini adalah kecepatan eksekusi yang tinggi, prediktabilitas hasil, dan kontrol penuh atas logika pencocokan, menjadikannya pilihan yang tepat untuk sistem yang memerlukan respons cepat dan deterministic based pada data internal perusahaan.

II.2.4 *Confidence Score dalam Intent Classification*

Confidence score atau skor keyakinan adalah nilai numerik antara 0 dan 1 yang menunjukkan tingkat kepercayaan sistem terhadap prediksi klasifikasi *intent* (Khosla dkk. 2022). Skor ini dihitung berdasarkan tingkat kecocokan antara input pengguna dengan pola yang telah didefinisikan. Semakin tinggi kecocokan, semakin tinggi nilai *confidence score*, dan semakin besar kemungkinan respons yang diberikan adalah benar.

Sistem menghitung *confidence score* dengan membandingkan input pengguna terhadap semua pola *utterances* atau frasa pelatihan yang tersedia. Sistem kemudian memilih *intent* dengan skor tertinggi sebagai prediksi. Implementasi *confidence score* memungkinkan sistem untuk mengidentifikasi kapan prediksi mungkin tidak

akurat dan memerlukan mekanisme *fallback* (Kuligowska dan Kowalczyk 2024).

II.2.4.1 *Confidence Threshold*

Confidence threshold adalah nilai ambang batas minimum yang harus dicapai oleh *confidence score* agar sistem memicu respons *intent* tertentu. Nilai threshold adalah angka numerik antara 0 dan 1, dengan nilai default yang umum digunakan adalah 0,45 atau 45% (Teneo Developers 2022). Sistem yang lebih konservatif dapat menggunakan threshold lebih tinggi seperti 0,70 atau 70% untuk mengurangi risiko respons yang salah.

Jika *confidence score* terbaik jatuh di bawah threshold, sistem akan memicu interaksi *fallback*, biasanya berupa pesan yang meminta klarifikasi atau memberitahu pengguna bahwa sistem tidak memahami pertanyaan. Pengaturan threshold yang tepat sangat penting untuk menyeimbangkan antara tingkat otomasi dan akurasi respons (Cyara 2024).

Analisis threshold membantu menentukan pengaturan yang optimal berdasarkan *Key Performance Indicators* (KPI) bisnis. Organisasi dapat menyesuaikan threshold berdasarkan tiga pendekatan utama. Pendekatan pertama adalah menyesuaikan threshold untuk mencapai persentase target respons yang benar. Pendekatan kedua adalah membatasi persentase maksimal respons yang salah. Pendekatan ketiga adalah mengoptimalkan trade-off antara tingkat otomasi dan risiko kesalahan berdasarkan toleransi bisnis (Cyara 2024; Genesys 2023).

II.2.5 *Evaluasi Performa Intent Classification*

Evaluasi performa klasifikasi *intent* dilakukan dengan mengukur beberapa metrik kunci yang mencerminkan akurasi dan keandalan sistem dalam memprediksi niat pengguna. Evaluasi yang komprehensif memastikan bahwa sistem dapat beroperasi sesuai dengan standar yang ditetapkan dan memberikan pengalaman pengguna yang optimal.

II.2.5.1 *Metrik Evaluasi Klasifikasi Intent Multi-Kelas*

Model klasifikasi intent dalam penelitian ini dievaluasi menggunakan metrik standar yang disesuaikan untuk masalah klasifikasi multi-kelas (38 kelas intent yang saling eksklusif), bukan multi-label. Pendekatan evaluasi harus mengakomodasi bahwa setiap kueri pengguna hanya dapat diklasifikasikan ke dalam satu kelas yang benar,

sehingga metrik klasifikasi biner harus diperluas menggunakan strategi *one-vs-all* dan agregasi yang tepat (Grandini, Bagli, dan Visani 2020; Evidently AI 2025).

Perbedaan utama antara klasifikasi multi-kelas dan biner adalah pada interpretasi elemen matriks kebingungan (*confusion matrix*) dan agregasi metrik per-kelas. Untuk 38 kelas, matriks kebingungan berukuran 38×38 , dengan setiap baris merepresentasikan kelas hasil prediksi dan setiap kolom merepresentasikan kelas sebenarnya. Prediksi yang benar terletak pada diagonal utama, sedangkan prediksi salah tersebar pada sel off-diagonal. Pola sebaran kesalahan pada matriks kebingungan multi-kelas memberikan wawasan berharga tentang persamaan atau kebingungan antar kelas, misalnya prediksi kelas “Penjualan-Bulanan” yang sering disalahklasifikasikan sebagai “Penjualan-Triwulanan” menunjukkan kesamaan semantik yang perlu ditangani dengan aturan atau templat yang lebih jelas (Evidently AI 2025).

Dalam menghitung metrik per kelas- i , masalah klasifikasi multi-kelas diubah menjadi masalah biner secara virtual melalui pendekatan *one-vs-all*: (Grandini, Bagli, dan Visani 2020)

- **TP_i** (*True Positive*): Jumlah sampel kelas- i yang diprediksi benar sebagai kelas- i (elemen diagonal matriks kebingungan pada baris dan kolom i)
- **FP_i** (*False Positive*): Jumlah sampel kelas selain i yang diprediksi sebagai kelas- i (jumlah seluruh baris i pada matriks kebingungan kecuali elemen diagonal)
- **FN_i** (*False Negative*): Jumlah sampel kelas- i yang diprediksi salah sebagai kelas lain (jumlah seluruh kolom i pada matriks kebingungan kecuali elemen diagonal)

Dalam konteks Agregasi Metrik (*Macro-Average* dengan *Micro-Average*) dan adanya 38 kelas dalam sistem ini, diperlukan strategi agregasi untuk memperoleh metrik keseluruhan performa sistem. Dua pendekatan utama yang umum digunakan dalam klasifikasi multi-kelas adalah *macro-averaging* dan *micro-averaging* (Evidently AI 2025)

Pendekatan ***Micro-Averaging*** ini menghitung metrik (*Precision*, *Recall*, *F1-Score*) untuk setiap kelas secara individual, kemudian mengambil rata-rata aritmatik dari 38 nilai metrik tersebut. Strategi ini memberikan bobot yang sama untuk setiap kelas, sehingga mengutamakan performa pada kelas minoritas dan mencegah kelas mayoritas mendominasi hasil evaluasi (Grandini, Bagli, dan Visani 2020).

$$\text{Presisi}_{\text{makro}} = \frac{1}{38} \sum_{i=1}^{38} \text{Presisi}_i \quad (\text{II.1})$$

$$\text{Daya Ingat}_{\text{makro}} = \frac{1}{38} \sum_{i=1}^{38} \text{Daya Ingat}_i \quad (\text{II.2})$$

$$\text{Skor-F1}_{\text{makro}} = \frac{1}{38} \sum_{i=1}^{38} \text{Skor-F1}_i \quad (\text{II.3})$$

Pendekatan **Micro-Averaging** menjumlahkan seluruh nilai TP, FP, dan FN dari semua kelas terlebih dahulu, kemudian menghitung metrik secara global berdasarkan jumlahan tersebut. Strategi ini memberikan bobot yang sama untuk setiap sampel, sehingga metrik cenderung didominasi oleh kelas mayoritas dan mengabaikan performa pada kelas minoritas. Untuk klasifikasi multi-kelas, mikro-rata presisi = mikro-rata daya ingat = akurasi (Evidently AI 2025).

$$\text{Presisi}_{\text{mikro}} = \frac{\sum_{i=1}^{38} \text{TP}_i}{\sum_{i=1}^{38} (\text{TP}_i + \text{FP}_i)} \quad (\text{II.4})$$

$$\text{Daya Ingat}_{\text{mikro}} = \frac{\sum_{i=1}^{38} \text{TP}_i}{\sum_{i=1}^{38} (\text{TP}_i + \text{FN}_i)} \quad (\text{II.5})$$

Sebagai pendekatan komplementer, rata-rata tertimbang menghitung metrik per-kelas kemudian mengambil rata-rata yang diboboti oleh jumlah sampel di setiap kelas atau **Weighted-Averaging**. Ini mengakomodasi keseimbangan antara kepedulian terhadap kelas minoritas dan representasi data yang ada (Grandini, Bagli, dan Visani 2020).

Sistem menggunakan mekanisme pemberian skor keyakinan (*confidence scoring*) dalam rentang 0 hingga 1 untuk setiap prediksi klasifikasi intent. Tingkat keyakinan tertinggi menunjukkan kelas yang diprediksi oleh sistem. Dalam menentukan apakah prediksi dapat diterima atau perlu klarifikasi, ditetapkan *threshold* pada nilai 0,70 (Santosa, Nusantara, dan Imron 2025):

- Jika tingkat keyakinan $\geq 0,70$: Prediksi dianggap “positif” dan diterima untuk kelas yang diprediksi; sistem memberikan respons berdasarkan intent yang teridentifikasi
- Jika tingkat keyakinan $< 0,70$: Prediksi dianggap “tidak cukup yakin”; sistem

meminta klarifikasi kepada pengguna atau menawarkan alternatif intent yang mungkin

Dengan demikian, untuk setiap kelas- i , perhitungan TP, FP, dan FN disesuaikan sebagai berikut:

- **TP _{i}** : Prediksi kelas- i dengan tingkat keyakinan $\geq 0,70$ yang ternyata benar
- **FP _{i}** : Prediksi kelas- i dengan tingkat keyakinan $\geq 0,70$ yang ternyata salah
- **FN _{i}** : Sampel kelas- i yang sebenarnya tidak terprediksi dengan tingkat keyakinan $\geq 0,70$ (termasuk prediksi kelas lain atau tingkat keyakinan rendah yang ditolak sistem)

Visualisasi matriks kebingungan untuk 38 kelas memberikan informasi penting tentang pola kesalahan klasifikasi:(Evidently AI 2025) Diagonal utama menampilkan jumlah prediksi benar per kelas, sementara intensitas warna pada sel off-diagonal menunjukkan kelas-kelas mana yang sering dikacaukan satu sama lain. Sebagai contoh, jika ditemukan frekuensi tinggi kesalahan pada sel (Penjualan-Bulanan, Penjualan-Triwulanan), hal ini menunjukkan bahwa dua kelas tersebut memiliki kesamaan semantik yang tinggi dan memerlukan aturan diferensiasi yang lebih spesifik atau contoh pelatihan yang lebih banyak.

Setiap kelas intent memiliki kepentingan strategis yang berbeda dalam konteks bisnis sehingga penelitian ini menggunakan kombinasi metrik berikut (Grandini, Bagli, dan Visani 2020):

- **Skor-F1 Makro** sebagai metrik utama: Mengutamakan performa yang merata di semua 38 kelas, memastikan bahwa kelas-kelas minoritas tidak terabaikan
- **Skor-F1 Tertimbang** sebagai metrik sekunder: Mempertimbangkan distribusi frekuensi kelas dalam data percobaan, mencerminkan performa yang lebih realistis terhadap data produksi
- **Skor-F1 Mikro** sebagai metrik verifikasi: Harus sama dengan akurasi keseluruhan; digunakan untuk memvalidasi perhitungan
- **Matriks Kebingungan**: Untuk analisis kualitatif pola kesalahan dan identifikasi peningkatan yang diperlukan pada aturan atau templat tertentu

Berdasarkan studi industri dan best-practices untuk sistem klasifikasi intent dalam chatbot bisnis, target performa yang ditetapkan untuk penelitian ini adalah:(Santosa, Nusantara, dan Imron 2025; Ozerova 2022)

- **Skor-F1 Makro** $\geq 0,85$: Menjamin performa yang baik dan seimbang di seluruh kelas, termasuk kelas-kelas dengan sampel pelatihan yang lebih sedikit
- **Skor-F1 Tertimbang** $\geq 0,90$: Menjamin performa yang baik pada kelas-

kelas mayoritas yang mewakili mayoritas pertanyaan pengguna dalam praktik

- **Skor-F1 Mikro** (Akurasi Keseluruhan) $\geq 0,88$: Target akurasi minimal untuk sistem dapat diterima dalam produksi
- **Matriks Kebingungan**: Diharapkan menunjukkan lebih dari 80% nilai terprediksi berada pada diagonal utama, mengindikasikan bahwa mayoritas prediksi benar

Pencapaian target-target ini akan diverifikasi melalui pengujian pada data validasi yang terpisah dan independen dari data pelatihan model, memastikan bahwa hasil evaluasi tidak bias dan representatif terhadap performa sistem pada data baru di lingkungan produksi.

II.2.5.2 *Confusion Matrix dan Confidence Score Analysis*

Confusion matrix menyediakan visualisasi detail tentang performa klasifikasi untuk setiap kelas *intent*. Analisis *confusion matrix* memungkinkan identifikasi *intent* yang sering salah diklasifikasikan dan pola kesalahan sistematis yang dapat diperbaiki melalui penyesuaian pola atau threshold.

Analisis distribusi *confidence score* penting untuk memahami bagaimana sistem membedakan antara *intent* yang benar (in-domain) dan pertanyaan di luar domain (out-of-domain atau OOD). Sistem yang baik harus menghasilkan *confidence score* tinggi untuk *intent* in-domain dan score rendah untuk OOD (Khosla dkk. 2022; Zhang dkk. 2022). Masalah *overconfidence*, dengan sistem menghasilkan score tinggi bahkan untuk sampel OOD yang abnormal, perlu dimitigasi melalui kalibrasi model yang tepat.

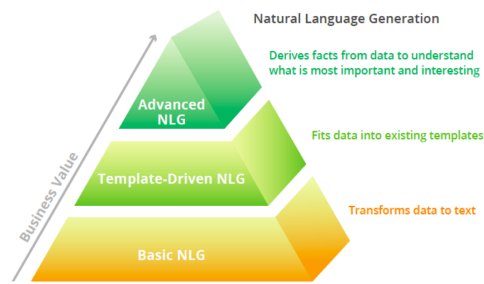
II.2.5.3 *Target Performa Intent Classification*

Berdasarkan literatur dan praktik industri terkini, target performa yang ditetapkan untuk sistem klasifikasi *intent* berbasis pola mencakup beberapa metrik kunci (Spot Intelligence 2023; Nabiilah, Adiwijaya, dan Maharani 2025). *Accuracy* harus mencapai minimal 85%, yang berarti minimal 85% dari semua prediksi *intent* harus benar. *F1-Score* ditargetkan minimal 0,80 untuk memastikan keseimbangan antara presisi dan *recall*. *Precision* ditetapkan minimal 0,82, yang berarti minimal 82% dari *intent* yang diprediksi positif benar-benar positif. *Recall* ditargetkan minimal 0,78, yang berarti minimal 78% dari *intent* positif aktual berhasil terdeteksi.

Metrik tambahan terkait *confidence score* meliputi *Average Confidence Score* untuk prediksi benar minimal 0,75, dan *Coverage Rate* atau persentase kueri yang menca-

pai threshold minimal 80%. *Response Time* harus di bawah 100 milidetik untuk 95% kueri agar memberikan pengalaman pengguna yang responsif. Pencapaian target-target ini memastikan bahwa sistem klasifikasi *intent* dapat memberikan pengalaman pengguna yang akurat dan responsif, dengan tingkat kesalahan yang minimal dan kemampuan untuk menangani variasi pertanyaan pengguna yang luas.

II.3 *Template-Based Natural Language Generation*



Gambar II.2 *Natural Language Generation*

Bagian ini membahas pendekatan pembangkitan bahasa alami berbasis templat untuk menghasilkan variasi respons dari data internal perusahaan. Pembahasan mencakup konsep NLG berbasis templat, keamanan data dalam implementasi NLG internal, dan evaluasi kualitas teks yang dihasilkan. Pendekatan ini memastikan bahwa respons yang dihasilkan bervariasi dan natural sambil tetap menjaga keamanan dan privasi data perusahaan.

II.3.1 *Konsep Natural Language Generation*

Natural Language Generation (NLG) adalah proses menghasilkan teks bahasa alami dari representasi data terstruktur. Dalam konteks *chatbot* Business Intelligence, NLG digunakan untuk mengubah hasil kueri basis data menjadi respons bahasa natural yang mudah dipahami pengguna (Kale dan Rastogi 2020). NLG memungkinkan sistem untuk menyajikan informasi numerik dan data terstruktur dalam format naratif yang lebih mudah dicerna.

Pasar NLP global diproyeksikan tumbuh dari USD 35,43 miliar pada tahun 2024 menjadi USD 438,08 miliar pada tahun 2034, dengan CAGR sebesar 28,6% (Precedence Research 2025). Pertumbuhan eksponensial ini didorong oleh meningkatnya adopsi teknologi AI, *cloud computing*, dan kebutuhan akan analisis data tidak terstruktur di berbagai sektor industri. Pertumbuhan pasar ini menunjukkan relevansi

dan urgensi pengembangan sistem NLG untuk mendukung kebutuhan bisnis modern.

II.3.2 *Template-Based Natural Language Generation*

Sistem NLG berbasis templat memetakan input non-linguistik secara langsung ke struktur linguistik yang berisi "celah" atau *placeholders* yang diisi selama output (Deemter, Krahmer, dan Theune 2005). Pendekatan ini menggunakan templat yang telah ditentukan sebelumnya dengan slot yang dapat diisi dengan data aktual dari hasil kueri basis data.

Contoh implementasi templat sederhana mencakup struktur seperti "Penjualan pada [bulan] mencapai [nilai], meningkat [persentase] % dari bulan sebelumnya." Sistem kemudian mengisi placeholder dengan data aktual: "Penjualan pada Oktober 2025 mencapai Rp 5,2 miliar, meningkat 15% dari bulan sebelumnya." (Kale dan Rastogi 2020).

Keuntungan utama pendekatan berbasis templat adalah kontrol penuh atas output, kecepatan generasi yang tinggi, tidak memerlukan dataset pelatihan besar, dan kemudahan maintenance dan modifikasi. Sistem dapat menghasilkan variasi output dengan mendefinisikan multiple templat untuk skenario yang sama, menghindari respons yang monoton (Kapoor dan Shetty 2023).

II.3.3 *Template Rewriting dengan Pre-trained Language Models*

Pendekatan lanjutan dalam NLG berbasis templat melibatkan penggunaan model bahasa pre-trained untuk menulis ulang templat sederhana menjadi teks yang lebih koheren dan natural. Metode ini menggabungkan keuntungan templat dengan fleksibilitas model bahasa neural (Kale dan Rastogi 2020; Rebuffel dkk. 2020).

Arsitektur sistem mencakup tiga tahapan utama. Tahap pertama adalah modul kebijakan menghasilkan sekumpulan tindakan berdasarkan konteks. Tahap kedua adalah templat sederhana mengkonversi setiap tindakan menjadi *utterance* bahasa alami. Tahap ketiga adalah model *encoder-decoder* seperti T5 menulis ulang gabungan *utterances* menjadi respons percakapan yang koheren (Kale dan Rastogi 2020).

Pendekatan ini memungkinkan sistem untuk menghasilkan respons yang secara semantik benar dari templat, kemudian model bahasa memperbaiki koherensi dan kealamian teks. Metode *template rewriting* telah menunjukkan peningkatan sample efficiency yang signifikan dibandingkan metode end-to-end, dengan jumlah templat

yang tumbuh linear terhadap jumlah slot dibandingkan pertumbuhan kombinatorial (Rebuffel dkk. 2020).

II.3.4 Keamanan dan Privasi Data dalam NLG Internal

Implementasi NLG dalam lingkungan perusahaan memerlukan pertimbangan keamanan yang ketat untuk melindungi data sensitif dan mencegah kebocoran informasi. Sistem NLG internal harus dirancang dengan prinsip *privacy-by-design* untuk memastikan data perusahaan tetap aman.

II.3.4.1 *Differential Privacy* untuk NLG

Differential Privacy (DP) adalah kerangka kerja yang memberikan jaminan privasi secara matematis, memastikan bahwa kontribusi individu dalam dataset dibatasi (Feyisetan dkk. 2022; Mireshghallah dkk. 2022). Dalam konteks NLG, DP dapat diterapkan melalui dua pendekatan utama. Pendekatan pertama adalah pelatihan model dengan DP-SGD yang menambahkan noise pada gradien selama pelatihan. Pendekatan kedua adalah perturbasi representasi teks yang menambahkan noise pada embedding sebelum generasi.

Parameter privasi ϵ mengontrol tingkat privasi yang diberikan, dengan nilai ϵ yang lebih kecil memberikan privasi yang lebih kuat tetapi dapat mengurangi utility model. Implementasi DP dalam NLG memerlukan trade-off yang cermat antara privasi dan kualitas output (Luo dan Liu 2024).

II.3.4.2 Strategi Keamanan Data dalam NLG

Implementasi NLG internal yang aman memerlukan beberapa strategi keamanan berlapis (IT Convergence 2024). Strategi pertama adalah enkripsi data dengan mengenkripsi data saat transit menggunakan TLS/SSL dan saat istirahat untuk melindungi data sensitif. Strategi kedua adalah kontrol akses yang ketat dengan menerapkan autentikasi dan otorisasi yang robust, menggunakan prinsip *least privilege*. Strategi ketiga adalah data minimization dengan hanya menggunakan data minimum yang diperlukan untuk generasi respons. Strategi keempat adalah *template-based approach* yang membatasi generasi hanya pada templat yang telah diverifikasi, mencegah generasi teks bebas yang dapat membocorkan informasi sensitif.

Strategi kelima adalah audit dan monitoring dengan mencatat semua aktivitas NLG untuk deteksi anomali dan audit keamanan. Strategi terakhir adalah *secure deplo-*

ymen environment dengan mengamankan infrastruktur server, firewall, dan menjaga software tetap up-to-date (IT Convergence 2024; Amazon Web Services 2024).

II.3.4.3 Mitigasi Risiko dalam NLG Internal

Risiko utama dalam implementasi NLG internal mencakup kebocoran data sensitif melalui output yang dihasilkan, *prompt injection* yang dapat memanipulasi model untuk menghasilkan output tidak diinginkan, dan *data tampering* yang dapat menyebabkan respons AI yang salah. Mitigasi risiko ini memerlukan pendekatan multi-layer (Fujitsu - PostgreSQL Fastware 2024).

Teknik mitigasi mencakup pemfilteran output untuk memeriksa dan menyensor informasi sensitif sebelum ditampilkan ke pengguna. *Input validation* digunakan untuk memvalidasi semua input dan menolak pola yang mencurigakan. *Template whitelisting* membatasi generasi hanya pada templat yang telah disetujui. Implementasi *role-based access control* (RBAC) memastikan pengguna hanya dapat mengakses data sesuai peran mereka. Regular security assessment dilakukan untuk menguji kerentanan sistem secara berkala (European Data Protection Board 2025; Velotix 2025).

II.3.5 Evaluasi Kualitas *Natural Language Generation*

Evaluasi kualitas output NLG merupakan aspek kritis dalam memastikan bahwa teks yang dihasilkan memenuhi standar kualitas dan sesuai dengan kebutuhan bisnis. Evaluasi dilakukan menggunakan berbagai metrik otomatis dan penilaian manusia.

II.3.5.1 Metrik Evaluasi NLG

Metrik evaluasi NLG dapat dikategorikan menjadi beberapa jenis berdasarkan karakteristiknya (Ni'mah dkk. 2023; Gehrmann dkk. 2022). Kategori pertama adalah metrik berbasis referensi yang membandingkan output dengan teks referensi yang ditulis manusia. Contoh metrik ini adalah BLEU yang mengukur kecocokan *n-gram* antara output dan referensi, dengan formula:

$$\text{BLEU} = BP \times \exp \left(\sum_{n=1}^N w_n \log p_n \right) \quad (\text{II.6})$$

dengan p_n adalah presisi *n-gram* dan BP adalah *brevity penalty*. Metrik lain adalah METEOR yang mempertimbangkan sinonim dan stemming, serta ROUGE yang

fokus pada *recall* untuk evaluasi summarization.

Kategori kedua adalah metrik berbasis embedding yang menggunakan representasi neural untuk mengukur kesamaan semantik. Contohnya adalah BERTScore yang menggunakan embedding BERT untuk menghitung kesamaan kontekstual antara tokens output dan referensi (Ni'mah dkk. 2023).

II.3.5.2 Metrik *Task-Specific* untuk BI Chatbot

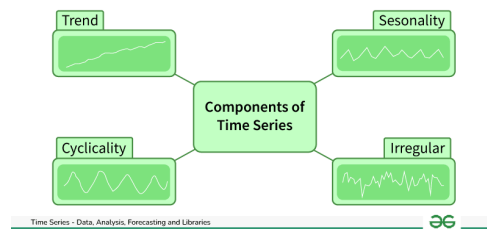
Evaluasi NLG dalam konteks *chatbot* BI memerlukan metrik tambahan yang spesifik untuk domain. Metrik pertama adalah *Factual Accuracy* yang mengukur kebenaran faktual dari informasi numerik dan data yang disajikan dalam respons. Metrik kedua adalah *Coherence* yang menilai koherensi dan alur logis dari respons yang dihasilkan. *Fluency* sebagai metrik ketiga mengukur kealamian dan kelancaran bahasa yang dihasilkan. Metrik keempat adalah *Relevance* yang menilai relevansi respons terhadap konteks kueri. Metrik terakhir adalah *Variability* yang mengukur keragaman respons untuk input yang sama atau serupa (Ni'mah dkk. 2023).

II.3.5.3 Target Performa NLG

Berdasarkan literatur dan praktik industri, target performa yang ditetapkan untuk sistem NLG berbasis templat dalam konteks *chatbot* BI mencakup beberapa kriteria utama (Ni'mah dkk. 2023; Gehrmann dkk. 2022). *Factual Accuracy* harus mencapai 100%, dengan tidak ada kesalahan dalam penyajian data numerik atau faktual. BLEU-4 score ditargetkan minimal 0,60 untuk memastikan kesamaan struktural yang baik dengan referensi natural. *Fluency* (penilaian manusia) harus mencapai minimal 4 dari skala 5, menunjukkan bahasa yang natural dan lancar. *Response Variety* ditargetkan minimal 5 variasi templat per kategori respons untuk menghindari monotonitas. *Generation Time* harus di bawah 50 milidetik untuk menjaga responsivitas sistem.

Pencapaian target-target ini memastikan bahwa sistem NLG dapat menghasilkan respons yang akurat, natural, dan bervariasi sambil tetap menjaga keamanan data internal perusahaan. Evaluasi berkelanjutan dan penyempurnaan templat diperlukan untuk mempertahankan kualitas output yang tinggi.

II.4 *Time Series Forecasting*



Gambar II.3 *Time Series Forecasting*

Bagian ini membahas konsep peramalan deret waktu, model-model statistik dan pembelajaran mendalam yang digunakan, serta evaluasi performa yang komprehensif. Pembahasan mencakup ARIMA, SARIMA, LSTM, serta perbandingan karakteristik dan performa masing-masing metode. Pemahaman yang mendalam tentang berbagai pendekatan peramalan ini menjadi fondasi penting dalam memilih dan mengimplementasikan model yang sesuai dengan karakteristik data dan kebutuhan aplikasi.

II.4.1 Konsep *Time Series* dan Peramalan

Time series adalah serangkaian titik data yang diindeks dalam urutan waktu, dan peramalan deret waktu adalah proses menggunakan model untuk memprediksi nilai masa depan berdasarkan nilai yang diamati sebelumnya. Peramalan akurat sangat penting untuk perencanaan strategis dan pengambilan keputusan di berbagai domain bisnis.

Pasar *time series forecasting* global mencapai USD 9,62 miliar pada tahun 2023 dan diperkirakan tumbuh menjadi USD 36,9 miliar pada tahun 2032 dengan CAGR sebesar 16,12% (WiseGuy Reports 2024). Pertumbuhan ini menunjukkan meningkatnya adopsi teknik peramalan deret waktu di berbagai industri untuk mendukung pengambilan keputusan berbasis data, termasuk peramalan permintaan layanan pelanggan, prediksi penjualan, dan analisis tren bisnis. Proyeksi pertumbuhan pasar ini mengindikasikan relevansi dan urgensi pengembangan sistem peramalan yang akurat dan dapat diandalkan.

II.4.2 Model ARIMA untuk *Time Series*

ARIMA (*AutoRegressive Integrated Moving Average*) adalah model statistik populer untuk peramalan deret waktu yang menggabungkan tiga komponen utama.

Model ARIMA dinyatakan sebagai $ARIMA(p, d, q)$, dengan p adalah jumlah *lag* observasi dalam model atau *order* dari komponen AR, d adalah jumlah kali deret waktu harus dibedakan untuk menjadi stasioner, dan q adalah ukuran jendela *moving average* atau *order* dari komponen MA.

Persamaan umum model ARIMA dapat ditulis sebagai:

$$\phi(B)(1 - B)^d X_t = \theta(B)\epsilon_t \quad (II.7)$$

Dalam persamaan ini, X_t adalah nilai deret waktu pada waktu t , B adalah operator *backshift*, $\phi(B)$ adalah polinomial autoregresif, $\theta(B)$ adalah polinomial *moving average*, ϵ_t adalah *white noise error term*, dan d adalah *order* pembedaan.

Model ARIMA efektif untuk data yang menunjukkan pola linear dan stasioner setelah pembedaan. Kelebihan utama ARIMA adalah interpretabilitas yang tinggi dan tidak memerlukan dataset besar untuk pelatihan, menjadikannya pilihan yang baik untuk aplikasi dengan data terbatas atau ketika transparansi model menjadi prioritas. Karakteristik ini membuat ARIMA tetap relevan meskipun telah muncul metode-metode yang lebih kompleks.

II.4.3 Model SARIMA untuk Data Musiman

SARIMA (*Seasonal ARIMA*) adalah ekstensi dari ARIMA yang dirancang untuk menangani data deret waktu musiman dengan lebih baik. SARIMA dinyatakan sebagai $SARIMA(p, d, q)(P, D, Q, m)$, dengan (p, d, q) adalah parameter non-musiman dan (P, D, Q) adalah parameter musiman, serta m sebagai jumlah langkah waktu untuk satu siklus musiman penuh.

Persamaan SARIMA dapat ditulis sebagai:

$$\phi(B)\Phi(B^m)(1 - B)^d(1 - B^m)^D X_t = \theta(B)\Theta(B^m)\epsilon_t \quad (II.8)$$

Dalam persamaan ini, $\Phi(B^m)$ dan $\Theta(B^m)$ adalah polinomial autoregresif dan *moving average* musiman.

SARIMA sangat cocok untuk data yang menunjukkan pola berulang pada interval waktu tetap, seperti penjualan bulanan yang cenderung meningkat pada bulan-bulan tertentu atau permintaan layanan pelanggan yang bervariasi mengikuti siklus musiman. Penambahan komponen musiman memungkinkan model untuk menangkap

dan memprediksi pola periodik yang tidak dapat ditangani oleh ARIMA standar. Kemampuan ini menjadikan SARIMA pilihan yang tepat untuk data bisnis yang umumnya menunjukkan pola musiman yang kuat.

II.4.4 Model LSTM untuk *Time Series*

LSTM (*Long Short-Term Memory*) adalah jenis jaringan saraf tiruan berulang yang dirancang khusus untuk mengatasi masalah *vanishing gradient* dalam RNN tradisional. LSTM mampu mempelajari dependensi jangka panjang dalam data sekuens, menjadikannya sangat efektif untuk peramalan deret waktu.

Arsitektur LSTM terdiri dari sel memori dan tiga gerbang utama. Gerbang lupa menentukan informasi apa dari sel memori yang harus dibuang. Gerbang masukan menentukan informasi baru apa yang akan disimpan dalam sel memori. Gerbang keluaran menentukan bagian mana dari sel memori yang akan menjadi keluaran.

Persamaan matematis LSTM dapat dituliskan sebagai berikut:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (\text{II.9})$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (\text{II.10})$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (\text{II.11})$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (\text{II.12})$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (\text{II.13})$$

$$h_t = o_t * \tanh(C_t) \quad (\text{II.14})$$

Dalam persamaan-persamaan ini, f_t adalah gerbang lupa, i_t adalah gerbang masukan, o_t adalah gerbang keluaran, C_t adalah keadaan sel, h_t adalah keadaan tersembunyi, σ adalah fungsi sigmoid, serta W dan b adalah parameter bobot dan *bias* yang dipelajari selama pelatihan.

LSTM unggul dalam menangani data deret waktu yang kompleks dengan dependensi jangka panjang dan pola non-linear. Kemampuan model untuk "mengingat" informasi relevan dari masa lalu yang jauh membuat LSTM sangat efektif untuk aplikasi seperti peramalan permintaan layanan pelanggan yang dipengaruhi oleh berbagai faktor temporal yang kompleks. Fleksibilitas dan kemampuan pembelajaran LSTM menjadikannya pilihan yang powerful untuk aplikasi dengan data yang kaya dan pola yang kompleks.

II.4.5 Perbandingan ARIMA, SARIMA, dan LSTM

Perbandingan komprehensif antara metode-metode peramalan deret waktu memberikan wawasan penting dalam memilih pendekatan yang sesuai untuk aplikasi spesifik. Pemahaman tentang kelebihan dan keterbatasan masing-masing metode memungkinkan pengambilan keputusan yang informed dalam pemilihan model.

II.4.5.1 Studi Empiris Perbandingan Model

Studi empiris menunjukkan bahwa algoritma berbasis *deep learning* seperti LSTM mengungguli algoritma tradisional seperti ARIMA (Siami-Namini, Tavakoli, dan Namin 2018). Rata-rata pengurangan tingkat kesalahan yang diperoleh oleh LSTM adalah antara 84-87% bila dibandingkan dengan ARIMA, menunjukkan keunggulan LSTM atas ARIMA dalam menangani data dengan pola kompleks dan non-linear.

Perbandingan ketiga model untuk prediksi profit menunjukkan akurasi sebesar 93,84% untuk ARIMA, 94,378% untuk SARIMA, dan 97,01% untuk LSTM (Sirisha, Belavagi, dan Attigeri 2022). LSTM memberikan akurasi tertinggi untuk prediksi profit. Hasil metrik kinerja menunjukkan bahwa ARIMA memiliki MAE = 33,05, RMSE = 39,28, dan $R^2 = 0,912$, SARIMA memiliki MAE = 31,48, RMSE = 38,06, dan $R^2 = 0,921$, sementara LSTM memiliki MAE = 28,63, RMSE = 34,91, dan $R^2 = 0,937$ (Sirisha, Belavagi, dan Attigeri 2022).

Perbandingan ARIMA, LSTM, dan GRU menunjukkan bahwa LSTM mengungguli model lain dengan MAPE sebesar 10,76%, diikuti oleh ARIMA pada 11,23% dan GRU pada 11,47% (Sunendar, Wibirama, dan Setiawan 2025). Temuan ini mengonfirmasi bahwa LSTM memiliki kemampuan generalisasi terbaik, terutama dalam menangani lompatan mendadak dan struktur temporal yang kompleks. Konsistensi hasil dari berbagai studi ini memperkuat posisi LSTM sebagai metode yang powerful untuk peramalan deret waktu yang kompleks.

II.4.5.2 Model Hibrida ARIMA-LSTM

Model hibrida ARIMA-LSTM menggabungkan keunggulan komplementer dari kedua pendekatan, menyediakan kerangka kerja yang lebih kuat dan akurat untuk peramalan deret waktu (Alharbi 2025). ARIMA, dikenal karena efektivitasnya dalam mengidentifikasi pola linear dan tren, membentuk fondasi model dengan menangani elemen stasioner dan linear dari deret waktu.

LSTM dirancang khusus untuk menangani dependensi non-linear dan hubungan

temporal jangka panjang dalam data. Dengan menggabungkan kedua pendekatan ini, model hibrida ARIMA-LSTM mendapatkan manfaat dari kemampuan pemodelan linear ARIMA dan pengenalan pola non-linear LSTM (Alharbi 2025). Integrasi ini memungkinkan model untuk menangkap dengan efektif baik fluktuasi jangka pendek maupun tren jangka panjang, menghasilkan prediksi yang lebih akurat dan robust terhadap berbagai kondisi data. Pendekatan hibrida ini memberikan solusi yang optimal dengan memanfaatkan kekuatan masing-masing metode.

II.4.5.3 Perbandingan Karakteristik Model

Studi komparatif menunjukkan bahwa LSTM lebih baik dalam menangkap pola musiman dan tren jangka panjang dibandingkan dengan ARIMA dan SARIMA untuk data konsumsi energi (Dubey, Kumar, dan Narang 2021). Tabel II.1 menunjukkan perbandingan karakteristik metode peramalan berdasarkan sintesis dari berbagai penelitian.

Tabel II.1 Perbandingan Metode Peramalan Deret Waktu

Metode	Kelebihan	Kekurangan	MAPE Tipikal
ARIMA	<ul style="list-style-type: none"> – <i>Interpretable</i> – Tidak memerlukan data besar – Cepat untuk dilatih 	<ul style="list-style-type: none"> – Sulit menangkap pola <i>non-linear</i> – Memerlukan data stasioner 	11–15%
SARIMA	<ul style="list-style-type: none"> – Menangani musiman – <i>Interpretable</i> – Baik untuk pola periodik 	<ul style="list-style-type: none"> – Memerlukan pengetahuan domain – Parameter banyak 	10–14%
LSTM	<ul style="list-style-type: none"> – Menangkap dependensi jangka panjang – Baik untuk pola kompleks – Tidak perlu stasioner 	<ul style="list-style-type: none"> – Memerlukan data besar – <i>Black box</i> – Lama untuk dilatih 	8–12%

Perbandingan ini menunjukkan bahwa pemilihan model harus disesuaikan dengan karakteristik data dan kebutuhan aplikasi spesifik. ARIMA dan SARIMA cocok untuk aplikasi yang memerlukan interpretabilitas tinggi dan memiliki data terbatas, sementara LSTM lebih sesuai untuk aplikasi dengan data yang berlimpah dan pola kompleks yang sulit dimodelkan dengan pendekatan statistik tradisional. Pemahaman tentang trade-off ini penting dalam membuat keputusan yang tepat untuk implementasi sistem peramalan.

II.4.6 Evaluasi Performa *Time Series Forecasting*

Evaluasi performa model peramalan deret waktu merupakan langkah kritis untuk memastikan bahwa prediksi yang dihasilkan memiliki akurasi yang memadai untuk mendukung pengambilan keputusan bisnis. Evaluasi yang komprehensif mencakup berbagai metrik dan teknik analisis yang saling melengkapi untuk memberikan gambaran menyeluruh tentang kualitas model.

II.4.6.1 Metrik Evaluasi *Forecasting*

Model peramalan deret waktu dievaluasi menggunakan tiga metrik utama yang masing-masing memberikan perspektif berbeda tentang akurasi prediksi (Sirisha, Belavagi, dan Attigeri 2022; Sunendar, Wibirama, dan Setiawan 2025). Metrik pertama adalah *Mean Absolute Error* (MAE) yang mengukur rata-rata dari nilai absolut kesalahan:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (\text{II.15})$$

MAE memberikan ukuran rata-rata kesalahan absolut tanpa mempertimbangkan arah kesalahan, baik positif maupun negatif. Metrik ini mudah diinterpretasikan dan robust terhadap *outliers*.

Metrik kedua adalah *Root Mean Square Error* (RMSE) yang mengukur akar kuadrat dari rata-rata kesalahan kuadrat:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (\text{II.16})$$

RMSE memberikan penalti yang lebih besar untuk kesalahan yang lebih besar karena proses kuadrat. Metrik ini lebih sensitif terhadap *outliers* dibandingkan MAE dan sering digunakan ketika kesalahan besar sangat tidak diinginkan.

Metrik ketiga adalah *Mean Absolute Percentage Error* (MAPE) yang mengukur rata-rata persentase kesalahan absolut:

$$\text{MAPE} = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (\text{II.17})$$

Dalam formula-formula ini, y_i adalah nilai aktual, \hat{y}_i adalah nilai prediksi, dan n

adalah jumlah observasi. MAPE memberikan ukuran kesalahan dalam bentuk persentase, memudahkan interpretasi dan perbandingan antar dataset dengan skala berbeda. Namun, MAPE tidak dapat digunakan ketika nilai aktual bernilai nol dan sangat sensitif terhadap nilai aktual yang kecil.

II.4.6.2 *Residual Analysis*

Analisis residual merupakan komponen penting dalam evaluasi model peramalan untuk memastikan bahwa asumsi model terpenuhi dan tidak ada pola sistematis yang tersisa dalam kesalahan prediksi. Analisis ini mencakup beberapa teknik yang saling melengkapi.

Teknik pertama adalah plot residual terhadap waktu untuk mendeteksi pola temporal yang mungkin tersisa. Residual yang baik harus terdistribusi secara acak di sekitar nol tanpa tren atau pola musiman yang jelas. Teknik kedua adalah uji normalitas residual menggunakan uji Shapiro-Wilk untuk menguji apakah residual mengikuti distribusi normal. Normalitas residual mengindikasikan bahwa model telah menangkap sebagian besar struktur dalam data. Teknik ketiga adalah uji autokorelasi residual menggunakan uji Ljung-Box untuk mendeteksi autokorelasi dalam residual. Autokorelasi yang signifikan mengindikasikan bahwa masih ada informasi temporal yang belum ditangkap oleh model. Ketiga teknik ini memberikan pemeriksaan komprehensif terhadap kualitas model dan asumsi yang mendasarinya.

II.4.6.3 *Cross-Validation untuk Time Series*

Validasi silang khusus untuk deret waktu (*time series cross-validation*) menggunakan pendekatan *walk-forward validation* atau *expanding window* untuk menghindari kebocoran informasi dari masa depan ke masa lalu. Pendekatan ini membagi data secara berurutan dan melatih model pada data historis untuk memprediksi periode mendatang, mensimulasikan kondisi *deployment* aktual dengan model harus memprediksi masa depan berdasarkan data masa lalu. Validasi silang deret waktu memberikan estimasi performa yang lebih realistis dibandingkan validasi silang standar yang dapat menghasilkan overfitting karena kebocoran informasi temporal.

II.4.6.4 *Target Performa Forecasting*

Berdasarkan literatur dan praktik industri terkini, target performa yang ditetapkan untuk model peramalan deret waktu dalam konteks permintaan layanan pelanggan mencakup beberapa kriteria utama (Siarni-Namini, Tavakoli, dan Namin 2018; Sunendar, Wibirama, dan Setiawan 2025; Sirisha, Belavagi, dan Attigeri 2022).

MAPE harus mencapai maksimal 15%, dengan kesalahan persentase rata-rata maksimal 15% yang dianggap dapat diterima untuk sebagian besar aplikasi bisnis. RMSE harus seminimal mungkin dan dibandingkan dengan standar deviasi data aktual, dengan RMSE yang kurang dari setengah standar deviasi data aktual umumnya dianggap baik. MAE harus seminimal mungkin dan diinterpretasikan dalam konteks skala data, dengan MAE yang secara signifikan lebih kecil dari mean nilai aktual menunjukkan performa yang baik.

Coefficient of Determination (R^2) harus mencapai minimal 0,85, yang berarti model harus dapat menjelaskan minimal 85% variansi dalam data. Residual harus lulus uji Ljung-Box dengan p-value lebih besar dari 0,05, mengindikasikan tidak ada autokorelasi signifikan yang tersisa. Pencapaian target-target ini memastikan bahwa model peramalan dapat memberikan prediksi yang akurat dan dapat diandalkan untuk mendukung perencanaan dan pengambilan keputusan strategis terkait permintaan layanan pelanggan.

II.5 Integrasi Sistem dan Penelitian Terkait

Bagian ini membahas integrasi berbagai komponen teknologi dalam sistem *chatbot* Business Intelligence, penelitian terkait arsitektur *chatbot* hibrida, serta aplikasi pembelajaran mesin untuk prediksi *customer churn*. Pembahasan mencakup strategi integrasi, arsitektur sistem, dan studi kasus implementasi yang menunjukkan efektivitas pendekatan terintegrasi. Pemahaman tentang integrasi sistem menjadi kunci dalam mengembangkan solusi yang komprehensif dan koheren.

II.5.1 Arsitektur Chatbot Hibrida

Arsitektur *chatbot* hibrida menggabungkan pendekatan berbasis aturan (*rule-based*) dengan kemampuan kecerdasan buatan untuk menciptakan sistem yang seimbang antara kontrol deterministik dan fleksibilitas adaptif. Pendekatan hibrida ini memungkinkan sistem untuk menangani berbagai jenis kueri dengan efisiensi optimal sambil mempertahankan akurasi dan keamanan data.

II.5.1.1 Komponen Arsitektur Chatbot Modern

Arsitektur *chatbot* modern terdiri dari lima komponen utama yang bekerja secara terintegrasi (Huang 2021; XenonStack 2023). Komponen pertama adalah antarmuka pengguna (*User Interface*) yang memungkinkan pengguna berkomunikasi dengan *chatbot* melalui platform perpesanan atau antarmuka berbasis web. Komponen ini

dapat mencakup konversi *speech-to-text* untuk input suara dan *text-to-speech* untuk output suara.

Komponen kedua adalah modul pemahaman bahasa alami (*Natural Language Understanding*) yang bertanggung jawab untuk menginterpretasikan masukan pengguna dan mengekstraksi *intent* serta entitas. Komponen ketiga adalah manajemen dialog (*Dialogue Management*) yang mengelola alur percakapan, memelihara konteks, dan menentukan tindakan yang sesuai berdasarkan *intent* yang teridentifikasi. Komponen keempat adalah *backend* yang terhubung dengan basis data, sistem eksternal, dan API untuk mengambil informasi yang diperlukan. Komponen terakhir adalah pembangkitan respons (*Response Generation*) yang menghasilkan output bahasa alami berdasarkan data yang diperoleh dan konteks percakapan (Huang 2021).

II.5.1.2 Chatbot Hibrida dengan AI dan Human Oversight

Pendekatan *chatbot* hibrida yang mengintegrasikan model NLP ringan dengan pengawasan manusia telah menunjukkan efektivitas tinggi dalam konteks pendidikan tinggi. Kerangka kerja hibrida ini menggunakan pohon keputusan, pemetaan *intent*, dan alur percakapan terstruktur untuk mengotomatisasi tugas rutin sambil mempertahankan dukungan kontekstual dan empatik melalui protokol eskalasi dinamis (Bamurange 2025).

Implementasi kerangka kerja hibrida pada Universitas Kigali menunjukkan akurasi 85% dalam menyelesaikan pertanyaan akademik dan mengurangi beban kerja staf sebesar 30% dalam simulasi (Bamurange 2025). Inovasi kunci mencakup fungsi *offline* untuk pengaturan dengan keterbatasan sumber daya, pemeriksaan kesesuaian budaya untuk menginterpretasikan kueri tidak langsung, dan strategi mitigasi bias yang selaras dengan pedoman etika. Temuan ini menekankan kelayakan platform *no-code* untuk deployment yang scalable, menekankan keseimbangan antara otomatisasi dan intervensi manusia.

II.5.1.3 Pasar Chatbot AI Global

Pasar *chatbot* berbasis AI global akan melebihi USD 3 miliar pada tahun 2025, dengan pasar *chatbot* hibrida suara dan teks mencapai USD 534 juta secara global pada tahun yang sama (Business Wire 2020). Penggunaan *Conversational AI* akan merevolusi manajemen hubungan pelanggan, dengan lebih dari 50% kueri pelanggan dapat dikelola melalui *chatbot* berbasis AI. Pertumbuhan eksponensial ini menunjukkan bahwa *chatbot* berbasis AI akan menjadi norma pada tahun 2025 ketika pusat

kontak meningkatkan otomasi.

Transformasi dari antarmuka operasional tradisional ke antarmuka percakapan akan mengubah secara dramatis ekspektasi tentang bagaimana manusia berkomunikasi, mengonsumsi konten, menggunakan aplikasi, dan terlibat dalam perdagangan (Business Wire 2020). Transformasi ini akan berdampak pada hampir setiap aspek operasi pemasaran dan penjualan untuk setiap vertikal industri, termasuk sektor Business Intelligence.

II.5.2 Integrasi *Pattern-Based Intent* dan *Template-Based NLG*

Integrasi antara klasifikasi *intent* berbasis pola dan pembangkitan bahasa alami berbasis templat menciptakan alur kerja yang efisien dan aman untuk sistem *chatbot* Business Intelligence. Integrasi ini memastikan bahwa kueri pengguna diproses dengan akurat sambil respons yang dihasilkan tetap bervariasi dan natural.

II.5.2.1 Alur Kerja Terintegrasi

Alur kerja sistem terintegrasi dimulai dengan penerimaan kueri pengguna melalui antarmuka *chatbot*. Modul klasifikasi *intent* berbasis pola melakukan pencocokan kueri dengan pola yang telah didefinisikan dan menghitung *confidence score*. Jika skor mencapai threshold yang ditentukan, sistem mengidentifikasi *intent* dan mengekstraksi parameter yang diperlukan untuk eksekusi kueri basis data.

Setelah *intent* teridentifikasi, sistem mengeksekusi kueri terhadap *data warehouse* untuk mengambil data internal yang relevan. Data yang diperoleh kemudian diteruskan ke modul NLG berbasis templat. Modul NLG memilih templat yang sesuai berdasarkan jenis *intent* dan konteks, mengisi *placeholder* dengan data aktual, dan menghasilkan respons bahasa alami. Untuk meningkatkan kealamian, sistem dapat menggunakan model *template rewriting* seperti T5 untuk memperbaiki koherensi dan variasi respons.

Pendekatan terintegrasi ini memastikan bahwa sistem dapat merespons dengan cepat untuk kueri terstruktur sambil menjaga keamanan data internal melalui pembatasan pada templat yang telah diverifikasi. Kontrol penuh atas output NLG mencegah kebocoran informasi sensitif dan memastikan bahwa respons tetap konsisten dengan kebijakan perusahaan.

II.5.2.2 Evaluasi Sistem Terintegrasi

Evaluasi sistem terintegrasi memerlukan pengukuran performa pada setiap komponen serta performa sistem secara keseluruhan. Metrik evaluasi mencakup akurasi klasifikasi *intent*, kualitas respons NLG, waktu respons end-to-end, dan kepuasan pengguna. Sistem yang baik harus mencapai akurasi klasifikasi *intent* minimal 85%, *factual accuracy* NLG 100%, dan waktu respons total di bawah 500 milidetik untuk 95% kueri.

Pengujian integrasi juga mencakup skenario pengujian end-to-end yang mensimulasikan percakapan nyata, pengujian beban untuk memastikan sistem dapat menangani volume kueri tinggi, dan pengujian keamanan untuk memverifikasi bahwa tidak ada kebocoran data sensitif. Pendekatan evaluasi holistik ini memastikan bahwa sistem tidak hanya berfungsi dengan baik pada tingkat komponen individual, tetapi juga memberikan pengalaman pengguna yang koheren dan memuaskan.

II.5.3 Customer Churn Prediction dengan Machine Learning

Prediksi *customer churn* adalah aplikasi pembelajaran mesin yang kritis dalam analisis bisnis pelanggan, memungkinkan organisasi untuk mengidentifikasi pelanggan berisiko dan mengambil tindakan retensi proaktif. Integrasi prediksi *churn* ke dalam sistem *chatbot* BI memberikan aksesibilitas yang mudah terhadap *insight* prediktif bagi pengguna internal.

II.5.3.1 Model Machine Learning untuk Churn Prediction

Beberapa model pembelajaran mesin telah terbukti efektif untuk prediksi *customer churn* (Pecan AI 2024; Kumar dan Zafar 2024). Model *Logistic Regression* dikenal karena kesederhanaannya dan kemudahan interpretasi, cocok untuk kasus dengan hubungan linear antara fitur dan *churn*. Model *Random Forest* memberikan akurasi superior dalam skenario kompleks dengan kemampuan menangani hubungan non-linear dan interaksi antar variabel. Model *Gradient Boosting Machines* (GBM) dan XGBoost sangat efektif dalam menangkap hubungan kompleks dan non-linear dengan membangun pohon secara sekuensial untuk memperbaiki kesalahan.

Studi terbaru menunjukkan bahwa XGBoost adalah algoritma paling efisien untuk prediksi *churn* pelanggan (Kumar dan Zafar 2024). Model hibrida berbasis jaringan saraf yang mengkombinasikan *Multi-Head Self-Attention*, BiLSTM, dan CNN telah diusulkan untuk meningkatkan ekstraksi fitur kompleks dan dependensi temporal (Li dkk. 2024). Model CCP-Net ini menggunakan algoritma sampling ADA-

SYN untuk menyeimbangkan ukuran sampel pelanggan yang *churn* dan tidak *churn*, mengatasi dampak negatif ketidakseimbangan sampel pada performa model.

II.5.3.2 Evaluasi Model *Churn Prediction*

Evaluasi model prediksi *churn* menggunakan metrik standar pembelajaran mesin dengan pertimbangan khusus untuk ketidakseimbangan kelas. Metrik evaluasi mencakup *Accuracy*, *Precision*, *Recall*, *F1-Score*, dan *ROC-AUC*. Studi Sharma, Gupta, dan Singh (2025) menunjukkan bahwa *Random Forest* mencapai akurasi 90,30% dengan semua atribut dan 90,90% dengan dataset atribut yang dikurangi, mengindikasikan bahwa dataset dengan atribut yang dikurangi dapat berguna untuk tugas prediksi *churn*.

Studi lain menunjukkan bahwa *Random Forest* mencapai akurasi 90,30%, presisi 89,5%, *recall* 91,2%, dan *ROC-AUC* 0,94 (Rahmat dan Setiawan 2024). Model dengan *hyperparameter tuning* menggunakan teknik seperti SMOTE, ENN, dan K-fold cross-validation menghasilkan peningkatan performa yang signifikan. Penggunaan dataset dengan atribut yang dikurangi tidak hanya meningkatkan akurasi tetapi juga mengurangi kompleksitas komputasi, menjadikannya lebih praktis untuk deployment produksi.

II.5.3.3 Fitur-fitur untuk *Churn Prediction*

Fitur-fitur yang umum digunakan untuk prediksi *churn* mencakup karakteristik demografis dan behavioral pelanggan (Pecan AI 2024; Sharma, Gupta, dan Singh 2025). Karakteristik demografis meliputi usia, jenis kelamin, lokasi geografis, dan tingkat pendapatan. Karakteristik behavioral mencakup *tenure* atau lama berlangganan, frekuensi transaksi, nilai transaksi rata-rata, variasi produk yang dibeli, dan pola penggunaan layanan.

Fitur tambahan yang relevan adalah riwayat keluhan atau masalah layanan, tingkat keterlibatan dengan program loyalitas, responsivitas terhadap kampanye pemasaran, dan perubahan pola penggunaan dari waktu ke waktu. *Feature engineering* yang efektif, termasuk pembuatan fitur agregat temporal dan fitur interaksi, dapat meningkatkan performa model secara signifikan.

II.5.3.4 *Real-Time Churn Prediction*

Pendekatan tradisional prediksi *churn* menggunakan pelatihan berbasis *batch* dengan dataset tetap yang dikumpulkan pada interval periodik. Pendekatan ini gagal

menangkap sifat dinamis industri dengan preferensi pelanggan yang berubah cepat (Al-Amri, Murugesan, dkk. 2025). Model RCE (*Real-time Continual learning-based Enhancement*) telah diusulkan untuk mengatasi keterbatasan ini dengan memanfaatkan pendekatan pengembangan *event-driven* untuk prediksi *churn* waktu nyata.

RCE menggunakan mekanisme *continual learning* berbasis *replay*, memungkinkan beradaptasi dengan perilaku pelanggan baru sambil mengurangi efek *catastrophic forgetting*. Berbeda dengan model berbasis *batch* tradisional, RCE memproses aliran data yang berkelanjutan, memungkinkan bisnis untuk bereaksi dengan cepat terhadap perubahan pasar (Al-Amri, Murugesan, dkk. 2025). Pendekatan ini memastikan bahwa pembuat keputusan menerima *insight* terkini ketika perilaku pelanggan berkembang, meningkatkan efektivitas strategi retensi.

II.5.4 Integrasi *Churn Prediction* dengan *Chatbot BI*

Integrasi model prediksi *churn* ke dalam sistem *chatbot* Business Intelligence memungkinkan pengguna internal untuk dengan mudah mengakses analisis risiko *churn* melalui antarmuka percakapan yang intuitif. Pengguna dapat mengajukan kueri seperti "Tampilkan pelanggan berisiko tinggi *churn* bulan ini" atau "Analisis faktor utama yang mempengaruhi *churn* di segmen premium", dan sistem akan mengeksekusi model prediksi serta menyajikan hasil dalam format bahasa alami yang mudah dipahami.

Integrasi ini memfasilitasi pengambilan keputusan proaktif untuk retensi pelanggan, seperti penawaran promosi khusus atau peningkatan layanan untuk pelanggan berisiko tinggi. Sistem dapat menghasilkan rekomendasi tindakan berdasarkan faktor-faktor yang berkontribusi terhadap risiko *churn* individu pelanggan. *Dashboard* interaktif dapat menampilkan distribusi risiko *churn* berdasarkan segmen pelanggan, tren *churn* dari waktu ke waktu, dan efektivitas kampanye retensi yang telah dilaksanakan.

Aksesibilitas informasi prediktif melalui antarmuka percakapan menurunkan *barrier* untuk adopsi analitik prediktif di kalangan pengguna non-teknis. Pengguna bisnis tanpa latar belakang *data science* dapat memanfaatkan model pembelajaran mesin yang canggih untuk mendukung keputusan mereka tanpa perlu memahami detail teknis implementasi model. Demokratisasi akses terhadap *insight* prediktif ini meningkatkan *data literacy* organisasi dan mempercepat transformasi menuju organisasi *data-driven*.

BAB III

ANALISIS MASALAH

Bab ini menyajikan analisis mendalam terhadap kondisi sistem *Business Intelligence* saat ini, identifikasi masalah pengguna, kebutuhan fungsional dan nonfungsional sistem, serta analisis pemilihan solusi teknologi yang paling tepat. Pembahasan mencakup evaluasi sistem *Business Intelligence* yang ada, identifikasi tantangan dalam pengambilan keputusan berbasis data, dan justifikasi pemilihan pendekatan *chatbot* berbasis aturan kueri, pemrosesan bahasa alami, dan peramalan deret waktu sebagai solusi yang paling sesuai dengan kebutuhan organisasi modern.

III.1 Analisis Kondisi Saat Ini

Analisis permasalahan yang dikaji dalam penelitian ini menggunakan pendekatan CRISP-DM (*Cross-Industry Standard Process for Data Mining*). Metodologi ini terdiri dari enam tahapan utama, yaitu: pemahaman bisnis (*business understanding*), pemahaman data (*data understanding*), penyiapan data (*data preparation*), pemodelan (*modeling*), evaluasi (*evaluation*), dan penerapan (*deployment*). Masing-masing tahapan dijelaskan lebih lanjut dalam bab ini beserta kegiatan-kegiatan yang dilakukan pada setiap fase, untuk menunjukkan bagaimana sistem dikembangkan secara terstruktur guna menjawab kebutuhan analisis dan prediksi kinerja bisnis secara otomatis dan responsif.

Permasalahan besar yang diangkat dalam penelitian ini adalah ketidakefisienan dan kurangnya responsivitas dalam proses analisis data bisnis untuk pengambilan keputusan strategis. Banyak organisasi, terutama perusahaan skala menengah dan besar, masih mengandalkan sistem *business intelligence* berbasis dasbor visual statis dan pelaporan manual yang memerlukan intervensi tim teknologi informasi atau analis data untuk setiap permintaan analisis berdasarkan kebutuhan khusus. Proses ini tidak hanya memakan waktu, tetapi juga berpotensi menimbulkan keterlambatan

dalam pengambilan keputusan yang berdampak pada daya saing perusahaan dan kemampuan merespons dinamika bisnis secara proaktif.

Kekurangan dari sistem *business intelligence* konvensional ini sebagian besar bersumber dari beberapa faktor kritis. Pertama, keterbatasan akses analitik berbasis percakapan yang menyebabkan pengguna internal tidak dapat memperoleh wawasan terkait indikator bisnis pelanggan seperti pesanan, pencapaian target penjualan, tingkat kepergian pelanggan, serta pendapatan secara cepat dan intuitif. Kedua, ketiadaan fitur prediksi yang mudah diakses oleh pengguna nonteknis, sehingga perencanaan kapasitas, evaluasi target, dan penetapan strategi bisnis menjadi kurang responsif terhadap tren dan dinamika masa depan. Ketiga, minimnya integrasi antara aturan kueri berbasis dan pemrosesan bahasa alami dalam sistem BI internal, yang menyebabkan proses pencarian data, penyaringan metrik utama, serta penyusunan laporan kinerja bisnis masih harus dilakukan secara manual dan berulang. Keempat, kurangnya analitik interaktif dan wawasan prediktif waktu nyata melalui antarmuka percakapan yang mudah diakses oleh seluruh pengguna internal.

Kondisi sistem saat ini dapat dijabarkan lebih rinci sebagai berikut:

1. **Responsivitas Terbatas**

Dasbor statis hanya menampilkan metrik yang telah didefinisikan sebelumnya dan tidak dapat menjawab pertanyaan mendadak atau spesifik dari pengguna bisnis. Pengguna harus menunggu tim teknologi informasi untuk membuat kueri khusus atau laporan baru, yang dapat memakan waktu beberapa jam hingga beberapa hari.

2. **Ketergantungan pada Tim Teknis**

Setiap permintaan kueri khusus atau analisis berdasarkan kebutuhan khusus harus diproses oleh tim teknologi informasi atau analis data, menciptakan kemacetan dan memperlambat waktu respons. Hal ini menghambat kemampuan organisasi untuk mengambil keputusan secara cepat dan tepat waktu.

3. **Kurangnya Kemampuan Prediktif**

Prediksi tren kinerja bisnis seperti volume penjualan, tingkat kepergian pelanggan, atau proyeksi pendapatan masih dilakukan secara manual menggunakan lembar sebar atau alat sederhana lainnya, tanpa memanfaatkan model peramalan yang canggih berbasis pembelajaran mesin atau peramalan deret waktu.

4. **Akses yang Tidak Fleksibel**

Sistem *business intelligence* umumnya hanya dapat diakses melalui komputer pribadi dengan antarmuka khusus yang memerlukan pengetahuan teknis,

tidak mendukung akses perangkat genggam atau antarmuka percakapan yang intuitif bagi pengguna nonteknis.

5. Integrasi Data Terbatas

Data berasal dari berbagai sumber seperti sistem manajemen hubungan pelanggan, sistem perencanaan sumber daya perusahaan, dan basis data transaksi, namun belum terintegrasi sepenuhnya dalam satu mesin kueri terpadu dan mudah diakses melalui antarmuka bahasa alami.

6. Ketiadaan Antarmuka Percakapan

Sistem *business intelligence* saat ini tidak menyediakan antarmuka percakapan yang memungkinkan pengguna berinteraksi dengan data menggunakan bahasa wajar, sehingga pengguna harus memahami struktur data dan terminologi teknis untuk dapat mengakses informasi yang dibutuhkan.

Permasalahan ini dapat diatasi melalui implementasi sistem *business intelligence* berbasis *chatbot* yang mengintegrasikan teknologi aturan kueri berbasis, pemrosesan bahasa alami, dan algoritma peramalan deret waktu. Pendekatan ini memungkinkan pengguna internal untuk mengajukan pertanyaan dalam bahasa wajar, mendapatkan respons yang cepat dan akurat, serta memperoleh prediksi tren bisnis secara otomatis tanpa memerlukan keahlian teknis khusus.

Pada penelitian ini, penulis memilih untuk mengimplementasikan klasifikasi maksud berbasis pencocokan pola dengan mekanisme nilai keyakinan untuk mengidentifikasi maksud pengguna dari kueri bahasa wajar, pembangkitan bahasa alami berbasis templat untuk menghasilkan respons naratif yang beragam namun tetap aman dan terkontrol, serta model peramalan deret waktu (ARIMA, SARIMA, atau LSTM) untuk menghasilkan prediksi terhadap indikator bisnis pelanggan seperti volume pesanan, pencapaian target penjualan, tingkat kepergian pelanggan, dan proyeksi pendapatan.

Oleh karena itu, berdasarkan rumusan masalah yang telah dijelaskan pada Bab I, penelitian ini bertujuan untuk merancang dan mengimplementasikan sebuah sistem *business intelligence* berbasis *chatbot* yang mampu memfasilitasi pengguna internal dalam menganalisis dan memprediksi indikator bisnis pelanggan secara interaktif dan waktu nyata. Sistem ini diharapkan dapat mengurangi beban kerja manual, meningkatkan responsivitas pengambilan keputusan, serta memberikan solusi terintegrasi yang siap diterapkan dalam praktik bisnis riil untuk meningkatkan daya saing organisasi di era digital.

III.2 Analisis Kebutuhan

Pada tahap awal pemahaman masalah dan solusi, analisis kebutuhan menjadi langkah dasar untuk memahami fitur dan karakteristik yang harus dimiliki oleh sistem agar dapat berfungsi sesuai tujuan. Dalam konteks tugas akhir ini, sistem dirancang untuk mengotomatisasi proses analisis dan prediksi kinerja bisnis berbasis data pelanggan menggunakan teknologi kecerdasan buatan, khususnya pemrosesan bahasa alami, klasifikasi maksud berbasis pola, dan peramalan deret waktu.

Analisis kebutuhan dibagi menjadi tiga kategori utama, yaitu identifikasi masalah pengguna, kebutuhan fungsional, dan kebutuhan nonfungsional. Identifikasi masalah pengguna berfokus pada pemahaman mendalam terhadap tantangan yang dihadapi oleh pengguna internal dalam mengakses dan menganalisis data bisnis. Kebutuhan fungsional berfokus pada fitur utama yang harus dimiliki sistem agar dapat menjalankan fungsinya secara efektif, seperti kemampuan memproses pertanyaan bahasa wajar, mengklasifikasikan maksud, mengeksekusi kueri ke gudang data, menghasilkan respons naratif, serta melakukan peramalan otomatis. Sementara itu, kebutuhan nonfungsional mencakup aspek teknis dan kualitas sistem, seperti keamanan data, performa, skalabilitas, ketersediaan, dan kemudahan penggunaan.

Pemahaman yang mendalam terhadap ketiga aspek kebutuhan ini menjadi fondasi penting dalam proses desain, implementasi, dan evaluasi sistem. Berikut ini adalah analisis kebutuhan yang telah disusun berdasarkan tujuan sistem dan karakteristik pengguna.

III.2.1 Identifikasi Masalah Pengguna

Pengguna dari sistem yang dikembangkan dalam penelitian ini adalah pengguna internal perusahaan, yang meliputi manajer, supervisor, analis bisnis, dan staf operasional yang terlibat dalam pengelolaan dan analisis data kinerja bisnis terkait pelanggan. Pengguna ini berasal dari berbagai departemen seperti penjualan, layanan pelanggan, pemasaran, dan manajemen strategis yang membutuhkan akses cepat terhadap wawasan bisnis untuk mendukung pengambilan keputusan harian maupun strategis.

Saat ini, mayoritas pengguna internal masih mengandalkan sistem *business intelligence* konvensional yang berbasis dasbor visual statis dan pelaporan manual. Pendekatan ini memiliki beberapa keterbatasan signifikan yang menghambat efektivitas pengambilan keputusan:

1. **Keterlambatan Respons terhadap Pertanyaan Bisnis**

Pengguna sering mengalami keterlambatan dalam memperoleh jawaban atas pertanyaan bisnis yang mendesak. Ketika membutuhkan wawasan spesifik yang tidak tersedia di dasbor yang ada, seperti perbandingan performa penjualan antar segmen pelanggan atau analisis tren kepergian pada periode tertentu, pengguna harus menunggu tim teknologi informasi atau analis data untuk membuat kueri khusus. Waktu tunggu ini dapat berkisar antara beberapa jam hingga beberapa hari, tergantung pada kompleksitas pertanyaan dan beban kerja tim teknis. Keterlambatan ini mengakibatkan pengambilan keputusan menjadi terhambat dan peluang bisnis yang peka waktu dapat terlewatkan.

2. **Kesulitan Mengakses Analisis Berdasarkan Kebutuhan Khusus**

Banyak pertanyaan bisnis yang bersifat berdasarkan kebutuhan khusus atau unik tidak dapat dijawab oleh dasbor statis yang telah dikonfigurasi sebelumnya. Pengguna harus menjelaskan kebutuhan analisis mereka secara detail kepada tim teknologi informasi melalui surat elektronik, sistem tiket, atau komunikasi langsung, yang kemudian memerlukan waktu bagi tim teknis untuk memahami persyaratan, membuat kueri yang sesuai, dan menghasilkan laporan. Proses ini tidak efisien, memakan waktu, dan rentan terhadap kesalahpahaman antara pengguna bisnis dan tim teknis yang dapat menyebabkan hasil analisis tidak sesuai dengan kebutuhan.

3. **Ketiadaan Kemampuan Prediksi Otomatis**

Analisis prediktif untuk tren kinerja bisnis seperti proyeksi volume penjualan, prediksi tingkat kepergian pelanggan, atau estimasi pencapaian target pendapatan masih dilakukan secara manual menggunakan lembar sebar atau alat analisis sederhana. Pengguna harus mengunduh data historis, melakukan perhitungan manual, atau meminta tim ilmu data untuk membuat prediksi khusus. Proses ini tidak dapat diperluas, memakan waktu, dan prediksi yang dihasilkan tidak selalu terkini mengikuti perubahan data terbaru. Ketiadaan prediksi otomatis membuat organisasi kurang siap dalam mengantisipasi tren atau risiko bisnis yang dapat berdampak pada strategi dan alokasi sumber daya.

4. **Akses yang Tidak Ramah Pengguna Nonteknis**

Dasbor *business intelligence* yang tersedia sering kali dirancang dengan asumsi bahwa pengguna memiliki pengetahuan teknis tentang struktur data, dimensi, metrik, dan terminologi basis data. Pengguna nonteknis, seperti manajer atau staf operasional yang tidak memiliki latar belakang teknis, sering merasa kesulitan dan kewalahan dengan berbagai tapis, daftar tarik, dan opsi yang ter-

sedia. Antarmuka tidak intuitif, tidak mendukung kueri berbasis bahasa wajar, dan memerlukan kurva pembelajaran yang tinggi. Hal ini menyebabkan banyak pengguna potensial tidak memanfaatkan sistem *business intelligence* secara optimal atau bahkan menghindari penggunaannya sama sekali.

5. Fragmentasi Akses Wawasan

Wawasan bisnis yang dibutuhkan untuk pengambilan keputusan tersebar di berbagai sistem, dasbor, laporan, dan alat yang berbeda. Pengguna harus membuka berbagai aplikasi atau modul untuk mendapatkan gambaran lengkap tentang kinerja bisnis, kemudian mengintegrasikan data secara manual di lembar sebar atau membuat dokumen analisis mereka sendiri. Fragmentasi ini meningkatkan risiko ketidakkonsistenan data, kesalahan manusia dalam proses konsolidasi, dan redundansi pekerjaan karena pengguna yang berbeda mungkin melakukan analisis yang sama secara terpisah tanpa berbagi hasil.

Tantangan-tantangan ini menghadirkan kebutuhan mendesak untuk mengembangkan sistem *business intelligence* yang lebih responsif, intuitif, dan terintegrasi, dengan kemampuan *chatbot* untuk antarmuka percakapan yang ramah pengguna, klasifikasi maksud yang cerdas untuk pengalihan kueri yang efisien, pembangkitan respons naratif yang beragam dan mudah dipahami, serta peramalan otomatis untuk prediksi kinerja bisnis. Sistem yang diusulkan diharapkan dapat mengatasi keterbatasan tersebut dan memberikan nilai tambah signifikan bagi organisasi dalam meningkatkan kecepatan, akurasi, dan kualitas pengambilan keputusan berbasis data.

III.2.2 Kebutuhan Fungsional

Kebutuhan fungsional menggambarkan fitur-fitur utama yang harus dimiliki oleh sistem *business intelligence* berbasis *chatbot* agar dapat memenuhi tujuan penelitian dan mengatasi masalah pengguna yang telah diidentifikasi. Setiap kebutuhan fungsional dirancang untuk mendukung proses analisis dan prediksi kinerja bisnis secara otomatis, responsif, dan mudah diakses oleh pengguna internal dengan berbagai tingkat keahlian teknis.

Berikut merupakan daftar kebutuhan fungsional yang telah disusun berdasarkan analisis mendalam terhadap proses bisnis, alur interaksi pengguna, dan arsitektur sistem yang diusulkan:

Tabel III.1 Kebutuhan Fungsional Sistem *Chatbot* BI

ID	Nama Kebutuhan	Deskripsi
FR-1	Penerimaan Pertanyaan Bahasa Alami	<p>Tujuan: Memungkinkan pengguna mengajukan pertanyaan atau permintaan analisis dalam bahasa Indonesia atau Inggris secara wajar melalui antarmuka <i>chatbot</i>.</p> <p>Masukan: Teks pertanyaan dari pengguna dalam bentuk bahasa wajar (contoh: "Berapa total penjualan bulan ini?" atau "Prediksi tingkat kepergian pelanggan 3 bulan ke depan").</p> <p>Keluaran: Pertanyaan tersimpan dan diproses oleh sistem untuk tahap berikutnya.</p>
FR-2	Klasifikasi Maksud dan Pemberian Nilai Keyakinan	<p>Tujuan: Mengidentifikasi maksud utama pengguna dari pertanyaan menggunakan pencocokan pola dan menghitung nilai keyakinan untuk menentukan tingkat kepercayaan sistem.</p> <p>Masukan: Teks pertanyaan hasil pemrosesan awal.</p> <p>Keluaran: Maksud teridentifikasi beserta nilai keyakinan (0-1) dan parameter terkait (entitas, dimensi waktu, metrik).</p>
FR-3	Pengalihan Kueri Berdasarkan Tingkat Keyakinan	<p>Tujuan: Mengarahkan kueri ke modul pemrosesan yang sesuai berdasarkan nilai keyakinan klasifikasi maksud.</p> <p>Masukan: Maksud dan nilai keyakinan.</p> <p>Keluaran: Jika nilai keyakinan $\geq 0,70$: kueri diteruskan ke mesin kueri berbasis aturan. Jika nilai keyakinan $< 0,50$: kueri diteruskan ke modul pemrosesan bahasa alami lanjutan atau permintaan klarifikasi ke pengguna.</p>

Bersambung ke halaman berikutnya

Tabel III.1 – Kebutuhan Fungsional Sistem *Chatbot* BI (lanjutan)

ID	Nama Kebutuhan	Deskripsi
FR-4	Eksekusi Kueri ke Gudang Data	<p>Tujuan: Mengambil data relevan dari gudang data atau mart data internal berdasarkan maksud dan parameter yang telah diidentifikasi.</p> <p>Masukan: Maksud, entitas, dan parameter kueri.</p> <p>Keluaran: Data hasil kueri dalam format terstruktur (JSON, tabel).</p>
FR-5	Penciptaan Respons Naratif Berbasis Templat	<p>Tujuan: Menghasilkan respons dalam bentuk teks naratif yang mudah dipahami menggunakan pembangkitan bahasa alami berbasis templat.</p> <p>Masukan: Data hasil kueri dan konteks percakapan.</p> <p>Keluaran: Teks respons naratif yang beragam dan wajar (contoh: "Penjualan pada bulan Oktober mencapai Rp 5,2 miliar, meningkat 15% dari bulan sebelumnya").</p>
FR-6	Dukungan Penggalian Data dan Analisis Perbandingan	<p>Tujuan: Memungkinkan pengguna melakukan eksplorasi data lebih dalam dari tingkat ringkasan ke tingkat detail, serta membandingkan metrik antar periode, segmen, atau dimensi.</p> <p>Masukan: Permintaan penggalian data atau perbandingan dari pengguna.</p> <p>Keluaran: Data detail atau hasil perbandingan dengan statistik perubahan (delta, persentase, tren).</p>
FR-7	Peramalan Deret Waktu	<p>Tujuan: Menghasilkan prediksi otomatis terhadap indikator kinerja bisnis menggunakan model ARIMA, SARIMA, atau LSTM.</p>

Bersambung ke halaman berikutnya

Tabel III.1 – Kebutuhan Fungsional Sistem *Chatbot* BI (lanjutan)

ID	Nama Kebutuhan	Deskripsi
		<p>Masukan: Data historis dan cakrawala prediksi yang diminta (harian, mingguan, bulanan).</p> <p>Keluaran: Nilai prediksi, selang kepercayaan, tren, dan visualisasi grafik.</p>
FR-8	Manajemen Riwayat Percakapan dan Sesi	<p>Tujuan: Menyimpan riwayat percakapan pengguna untuk memudahkan pelacakan dan konteks percakapan yang berkelanjutan.</p> <p>Masukan: Interaksi pengguna dengan sistem.</p> <p>Keluaran: Riwayat percakapan tersimpan dan dapat diakses kembali oleh pengguna.</p>
FR-9	Autentikasi dan Otorisasi Pengguna	<p>Tujuan: Memastikan hanya pengguna yang berwenang dapat mengakses sistem dan data sesuai dengan peran mereka.</p> <p>Masukan: Kredensial pengguna (nama pengguna, sandi).</p> <p>Keluaran: Akses sistem berdasarkan kendali akses berbasis peran.</p>
FR-10	Mekanisme Pengunduran dan Penanganan Kesalahan	<p>Tujuan: Menangani kueri yang tidak dapat dipahami atau diproses dengan memberikan respons alternatif atau permintaan klarifikasi.</p> <p>Masukan: Kueri dengan nilai keyakinan rendah atau kesalahan pemrosesan.</p> <p>Keluaran: Pesan klarifikasi, saran pertanyaan alternatif, atau eskalasi ke dukungan manusia.</p>

Kebutuhan-kebutuhan fungsional di atas dirancang untuk menciptakan sistem yang komprehensif, responsif, dan mudah digunakan, yang mampu mengotomatisasi sebagian besar tugas analisis dan prediksi bisnis, mengurangi ketergantungan pada tim teknis, serta memberdayakan pengguna internal untuk mengakses wawasan yang

mereka butuhkan dengan cepat dan efisien.

III.2.3 Kebutuhan Nonfungsional

Kebutuhan nonfungsional menggambarkan karakteristik kualitas sistem yang harus dipenuhi untuk memastikan sistem dapat beroperasi dengan baik, aman, andal, dan memberikan pengalaman pengguna yang optimal. Kebutuhan nonfungsional mencakup aspek-aspek seperti performa, keamanan, ketersediaan, skalabilitas, kemudahan penggunaan, dan pemeliharaan sistem.

Berikut merupakan daftar kebutuhan nonfungsional yang telah disusun untuk sistem *Business Intelligence* berbasis *chatbot*:

Tabel III.2 Kebutuhan Nonfungsional Sistem *Chatbot* BI

ID	Nama Kebutuhan	Deskripsi
NF-1	Keamanan Data dan Privasi	Sistem harus memastikan keamanan data internal perusahaan dengan menerapkan enkripsi data saat dalam perjalanan (TLS/SSL) dan saat penyimpanan. Sistem harus mengimplementasikan kendali akses berbasis peran yang ketat untuk memastikan pengguna hanya dapat mengakses data sesuai dengan peran dan wewenang mereka. Sistem harus mencegah kebocoran data sensitif, suntikan perintah berbahaya, atau akses tidak sah. Pencatatan dan jejak audit untuk semua aktivitas kueri dan akses data harus diimplementasikan untuk kepatuhan dan analisis forensik.

Bersambung ke halaman berikutnya

Tabel III.2 – Kebutuhan Nonfungsional Sistem *Chatbot* BI (lanjutan)

ID	Nama Kebutuhan	Deskripsi
NF-2	Ketersediaan Sistem	Sistem harus memiliki tingkat ketersediaan minimal 99% agar dapat diakses kapan saja oleh pengguna internal. Sistem harus dilengkapi dengan mekanisme redundansi dan pengalihan otomatis untuk memastikan kontinuitas layanan. Jika terjadi kesalahan pada saat eksekusi kueri atau penciptaan respons, sistem harus memiliki mekanisme pengunduran yang elegan dan penanganan kesalahan yang memadai serta memberikan pesan kesalahan yang informatif kepada pengguna.
NF-3	Performa dan Waktu Respons	Sistem harus mampu memberikan waktu respons maksimal 2 detik untuk 95% dari kueri tipikal. Untuk kueri kompleks yang memerlukan peramalan atau analisis data besar, waktu respons maksimal adalah 5 detik. Sistem harus dapat memproses minimal 100 kueri per menit pada kondisi beban normal dan hingga 1000 kueri per jam pada kondisi beban puncak tanpa penurunan performa yang signifikan.
NF-4	Skalabilitas	Sistem harus mampu menangani peningkatan jumlah pengguna dan volume data secara horizontal dengan menambahkan server atau simpul tambahan. Infrastruktur harus dirancang dengan arsitektur layanan mikro atau penerapan terkontainer untuk memudahkan penskalaan. Basis data dan mesin kueri harus dioptimalkan untuk menangani volume data yang besar tanpa penurunan performa.

Bersambung ke halaman berikutnya

Tabel III.2 – Kebutuhan Nonfungsional Sistem *Chatbot* BI (lanjutan)

ID	Nama Kebutuhan	Deskripsi
NF-5	Kemudahan Penggunaan	Antarmuka <i>chatbot</i> harus intuitif dan mudah digunakan oleh pengguna nonteknis tanpa memerlukan pelatihan ekstensif. Sistem harus menyediakan panduan kontekstual dan saran pertanyaan untuk membantu pengguna merumuskan kueri. Riwayat percakapan dan manajemen sesi harus didukung untuk memudahkan pengguna melacak interaksi sebelumnya. Mekanisme umpan balik harus tersedia untuk perbaikan berkelanjutan berdasarkan masukan pengguna.
NF-6	Kemudahan Pemeliharaan	Sistem harus mudah dipelihara dengan dokumentasi kode, antarmuka program aplikasi, dan arsitektur yang lengkap dan jelas. Templat, aturan berbasis pola, dan pemetaan maksud harus mudah diperbarui oleh tim tanpa memerlukan penerapan ulang sistem atau waktu henti. Sistem harus dirancang secara modular sehingga mudah untuk menambahkan maksud baru, aturan baru, atau model peramalan baru di masa depan.

Bersambung ke halaman berikutnya

Tabel III.2 – Kebutuhan Nonfungsional Sistem *Chatbot* BI (lanjutan)

ID	Nama Kebutuhan	Deskripsi
NF-7	Kompatibilitas dan Aksesibilitas	Sistem harus dapat diakses dari berbagai platform (<i>web</i> , perangkat genggam, komputer pribadi) dan perangkat (ponsel pintar, tablet, komputer pribadi/laptop) dengan faktor bentuk yang berbeda. Antarmuka harus responsif dan mendukung berbagai ukuran layar. Sistem harus kompatibel dengan peramban umum seperti Chrome, Firefox, Edge, dan Safari, serta mendukung berbagai sistem operasi (Windows, macOS, Linux, iOS, Android).
NF-8	Keandalan	Sistem harus memiliki tingkat keandalan yang tinggi dengan kemampuan menangani kesalahan dan kondisi tidak normal tanpa berhenti. Sistem harus memiliki mekanisme pengulangan otomatis untuk proses yang gagal dan pencatatan komprehensif untuk analisis penyebab akar. Hasil prediksi dan analisis harus konsisten dan dapat direproduksi dengan data masukan yang sama.
NF-9	Integritas Data	Sistem harus memastikan bahwa data yang diambil, diproses, dan disajikan kepada pengguna adalah akurat dan tidak mengalami kerusakan. Hasil ekstraksi dari gudang data tidak boleh rusak, hilang, atau berubah saat disimpan, diproses, atau diekspor. Sistem harus mengimplementasikan validasi data dan penjumlahan cek untuk memastikan integritas data di seluruh saluran pemrosesan.

Bersambung ke halaman berikutnya

Tabel III.2 – Kebutuhan Nonfungsional Sistem *Chatbot* BI (lanjutan)

ID	Nama Kebutuhan	Deskripsi
NF-10	Ekstensibilitas	Sistem harus dirancang dengan arsitektur yang fleksibel dan modular untuk memungkinkan penambahan fitur baru di masa depan, seperti integrasi dengan sistem eksternal (ERP, CRM), penambahan model pembelajaran mesin baru, atau perluasan ke domain bisnis lainnya. Antarmuka program aplikasi yang terdokumentasi dengan baik harus disediakan untuk memfasilitasi integrasi dengan sistem lain.

Kebutuhan nonfungsional ini dirancang untuk memastikan bahwa sistem tidak hanya dapat menjalankan fungsi-fungsi yang diharapkan, tetapi juga memiliki kualitas teknis yang tinggi, aman, andal, dan memberikan pengalaman pengguna yang memuaskan. Pencapaian kebutuhan nonfungsional ini akan dievaluasi melalui pengujian performa, pengujian keamanan, pengujian beban, dan survei kepuasan pengguna pada tahap implementasi dan evaluasi sistem.

III.3 Analisis Pemilihan Solusi

Analisis permasalahan yang dikaji dalam penelitian ini menggunakan pendekatan kesesuaian masalah-solusi, yaitu dengan mencocokkan kebutuhan nyata yang terjadi di lapangan dengan solusi teknologi yang paling tepat guna. Pendekatan ini mempertimbangkan berbagai alternatif solusi yang tersedia, mengevaluasi kelebihan dan kekurangan masing-masing, serta memilih solusi yang paling optimal berdasarkan kriteria yang telah ditetapkan menggunakan metode *Balanced Scorecard*.

Pemilihan solusi yang tepat sangat penting untuk memastikan bahwa sistem yang dikembangkan dapat mengatasi masalah pengguna secara efektif, memberikan nilai tambah yang signifikan bagi organisasi, serta berkelanjutan untuk pengembangan jangka panjang. Analisis ini akan membantu memvalidasi bahwa solusi yang dipilih adalah yang paling sesuai dengan kondisi, kebutuhan, dan tujuan strategis organisasi.

III.3.1 Alternatif Solusi

Dalam rangka meningkatkan efisiensi analisis dan prediksi kinerja bisnis serta mengatasi keterbatasan sistem *business intelligence* konvensional, terdapat beberapa alternatif solusi teknologi yang dapat dipertimbangkan. Setiap alternatif memiliki karakteristik, kelebihan, dan kekurangan yang berbeda, yang perlu dievaluasi secara komprehensif untuk menentukan solusi yang paling optimal.

Berikut adalah tiga alternatif solusi utama yang telah diidentifikasi dan dianalisis berdasarkan studi literatur, praktik industri, dan kebutuhan spesifik organisasi:

Tabel III.3 Perbandingan Alternatif Solusi Teknologi

No	Nama Solusi	Deskripsi	Kelebihan	Kekurangan
1	<i>Chatbot</i> Berbasis Aturan Sederhana	Menggunakan <i>chatbot</i> dengan pencocokan pola yang sangat kaku, hanya merespons pola kalimat yang telah ditentukan sebelumnya. Sistem hanya dapat menangani kueri terstruktur dengan format tetap, tidak mampu memahami variasi bahasa wajar yang kompleks. Tidak ada model peramalan, hanya pengambilan data dari basis data.	<ul style="list-style-type: none">• Implementasi cepat dan hemat biaya• Mudah dipelihara untuk aturan sederhana• Tidak memerlukan infrastruktur kompleks• Kontrol penuh atas respons sistem	<ul style="list-style-type: none">• Tidak dapat diperluas untuk kebutuhan BI modern• Pengguna harus menyesuaikan dengan pola tetap• Tidak mampu menjawab kueri kompleks• Tidak ada kemampuan prediktif• Pengalaman pengguna terbatas

Bersambung ke halaman berikutnya

Tabel III.3 – Perbandingan Alternatif Solusi Teknologi (lanjutan)

No	Nama Solusi	Deskripsi	Kelebihan	Kekurangan
2	<i>Chatbot</i> BI dengan Klasifikasi Maksud Berbasis Pola, Pembangkitan Bahasa Alami Berbasis Templat, dan Peramalan Deret Waktu	Mengimplementasikan sistem BI <i>chatbot</i> yang terintegrasi dengan tiga komponen: klasifikasi maksud berbasis pola dengan pemberian nilai keyakinan, pembangkitan bahasa alami berbasis templat untuk respons yang beragam namun aman, dan model peramalan deret waktu (ARIMA, SARIMA, LSTM) untuk prediksi otomatis.	<ul style="list-style-type: none"> • Mampu menangani variasi bahasa wajar kompleks • Respons beragam dan wajar • Dukungan peramalan otomatis • Adaptif terhadap kebutuhan dinamis • Pengalaman pengguna unggul • Keamanan terjaga dengan templat • Dapat diperluas secara fleksibel 	<ul style="list-style-type: none"> • Kompleksitas implementasi lebih tinggi • Memerlukan usaha awal untuk desain • Pemeliharaan butuh keahlian berbagai bidang • Biaya infrastruktur lebih tinggi

Bersambung ke halaman berikutnya

Tabel III.3 – Perbandingan Alternatif Solusi Teknologi (lanjutan)

No	Nama Solusi	Deskripsi	Kelebihan	Kekurangan
3	Dasbor BI Interaktif Tradisional dengan Peningkatan	Tetap menggunakan dasbor BI visual yang sudah ada, dengan peningkatan untuk interaktivitas yang lebih baik. Pengguna dapat melakukan penggalian data, penyaringan, dan agregasi melalui kontrol antarmuka pengguna yang lebih intuitif. Menambahkan fitur analitik layanan mandiri dan aplikasi dasbor perangkat genggam.	<ul style="list-style-type: none"> • Teknologi matang dan terbukti • Dukungan komunitas dan alat kaya • Tata kelola data dan keamanan mapan • Akrab bagi organisasi dengan alat BI tradisional 	<ul style="list-style-type: none"> • Tidak mendukung kueri bahasa wajar • Tidak ada antarmuka percapakan • Dukungan terbatas untuk peramalan • Tetap memerlukan keahlian teknis • Tidak dapat diperluas untuk bisnis peka waktu

Bersambung ke halaman berikutnya

Tabel III.3 – Perbandingan Alternatif Solusi Teknologi (lanjutan)

No	Nama Solusi	Deskripsi	Kelebihan	Kekurangan
4	Chatbot BI Berbasis Model Bahasa Besar (Large Language Model)	Mengimplementasikan chatbot BI yang memanfaatkan arsitektur Transformer penuh dengan model bahasa besar pra-latih seperti GPT, BERT, atau varian lainnya. Sistem menggunakan pembelajaran mendalam untuk pemahaman bahasa alami yang sangat canggih, mampu menangani kueri kompleks dengan konteks multiturn, dan menghasilkan respons naratif yang sangat wajar. Model dapat disetel-halus (fine-tuned) dengan data domain spesifik untuk meningkatkan akurasi.	<ul style="list-style-type: none"> • Pemahaman bahasa alami paling canggih • Kemampuan konteks percakapan mendalam • Respons sangat wajar dan beragam • Dapat menangani kueri sangat kompleks • Adaptif terhadap variasi bahasa ekstrem • Pembelajaran berkelanjutan dari interaksi 	<ul style="list-style-type: none"> • Biaya komputasi sangat tinggi • Memerlukan infrastruktur GPU berkinerja tinggi • Risiko halusinasi dan respons tidak akurat • Sulit mengontrol keluaran secara deterministik • Potensi kebocoran data sensitif • Kompleksitas implementasi sangat tinggi • Ketergantungan pada penyedia model eksternal • Latensi respons lebih tinggi

Alternatif 1: Chatbot Berbasis Aturan Sederhana. Solusi ini mengimplementasikan *chatbot* dengan pendekatan berbasis aturan yang sederhana, menggunakan pencocokan pola berbasis ekspresi reguler atau pencocokan langsung. Sistem hanya dapat merespons pertanyaan dengan format yang telah didefinisikan sebelumnya

dan tidak memiliki kemampuan untuk memahami variasi bahasa atau konteks yang kompleks. Fitur yang tersedia terbatas pada pengambilan data sederhana dari basis data tanpa kemampuan analisis lanjutan atau prediksi. Antarmuka juga terbatas dan tidak dioptimasi untuk pengalaman pengguna yang baik pada perangkat genggam.

Alternatif 2: *Chatbot* BI dengan Klasifikasi Maksud Berbasis Pola, Pembangkitan Bahasa Alami Berbasis Templat, dan Peramalan Deret Waktu. Solusi ini merupakan pendekatan komprehensif yang mengintegrasikan tiga komponen teknologi utama. Pertama, klasifikasi maksud berbasis pola dengan mekanisme pemberian nilai keyakinan untuk mengidentifikasi maksud pengguna dari kueri bahasa wajar dengan akurasi tinggi. Kedua, pembangkitan bahasa alami berbasis templat untuk menghasilkan respons naratif yang variatif, wajar, dan mudah dipahami sambil menjaga keamanan data internal melalui pembatasan pada templat yang telah diverifikasi. Ketiga, model peramalan deret waktu (ARIMA, SARIMA, atau LSTM) untuk menghasilkan prediksi otomatis terhadap indikator kinerja bisnis seperti volume penjualan, tingkat kepergian pelanggan, dan proyeksi pendapatan. Sistem ini dirancang dengan arsitektur modular yang dapat diperluas dan diperbesar, mendukung akses berbagai platform (web, perangkat genggam, komputer pribadi), dan menerapkan kontrol keamanan yang ketat.

Alternatif 3: Dasbor BI Interaktif Tradisional dengan Peningkatan. Solusi ini mempertahankan pendekatan dasbor BI visual yang sudah ada dengan menambahkan peningkatan pada aspek interaktivitas dan kemudahan penggunaan. Fitur penggalan data lebih dalam, penyaringan lanjutan, dan pembuatan laporan khusus ditingkatkan dengan antarmuka yang lebih intuitif. Sistem ini juga dilengkapi dengan aplikasi dasbor perangkat genggam untuk akses data dari perangkat bergerak. Namun, solusi ini tidak menyediakan antarmuka percakapan berbasis bahasa wajar dan kemampuan prediktif yang otomatis.

Alternatif 4: *Chatbot* BI Berbasis Model Bahasa Besar (*Large Language Model*). Solusi ini mengimplementasikan *chatbot* BI yang memanfaatkan arsitektur *Transformer* penuh dengan model bahasa besar pra-latih seperti GPT, BERT, T5, atau varian lainnya. Sistem menggunakan pembelajaran mendalam untuk pemahaman bahasa alami yang sangat canggih, mampu menangani kueri kompleks dengan konteks percakapan multiturn, dan menghasilkan respons naratif yang sangat wajar menyerupai komunikasi manusia. Model dapat disetel-halus (*fine-tuned*) dengan data domain spesifik organisasi untuk meningkatkan akurasi dan relevansi respons terhadap konteks bisnis. Sistem ini memanfaatkan kemampuan pemodelan ba-

hasa yang telah dilatih pada korpus teks masif untuk memahami nuansa, konteks implisit, dan melakukan inferensi kompleks dari pertanyaan pengguna. Namun, solusi ini memerlukan infrastruktur komputasi dengan GPU berkinerja tinggi untuk inferensi cepat, menimbulkan tantangan dalam hal kontrol deterministik terhadap keluaran, dan menghadapi risiko halusinasi di mana model dapat menghasilkan respons yang terdengar meyakinkan tetapi faktanya tidak akurat atau tidak berdasar pada data nyata. Implementasi juga memerlukan pertimbangan ketat terhadap keamanan data untuk mencegah kebocoran informasi sensitif melalui model, serta strategi mitigasi risiko untuk memastikan keandalan respons dalam konteks pengambilan keputusan bisnis kritis.

III.3.2 Analisis Penentuan Solusi

Dalam menentukan solusi terbaik untuk meningkatkan sistem *Business Intelligence* dan mengatasi masalah yang telah diidentifikasi, dilakukan analisis menggunakan metode *Balanced Scorecard* (BSC). Metode ini menilai masing-masing solusi dari empat perspektif utama yang saling melengkapi: *Financial*, *Customer*, *Internal Business Process*, dan *Learning & Growth*. Pendekatan holistik ini memastikan bahwa keputusan pemilihan solusi tidak hanya mempertimbangkan aspek biaya, tetapi juga nilai strategis jangka panjang, kepuasan pengguna, efisiensi proses bisnis, dan kapasitas inovasi organisasi.

Berikut adalah analisis mendalam untuk setiap solusi berdasarkan keempat perspektif *Balanced Scorecard*:

Solusi 1: Chatbot Berbasis Aturan Sederhana

Tabel III.4 Analisis *Balanced Scorecard* untuk Aturan Sederhana

Perspektif	Penilaian
<i>Financial</i>	Solusi ini memerlukan investasi awal yang rendah hingga sedang karena tidak membutuhkan infrastruktur komputasi yang kompleks atau model pembelajaran mesin yang canggih. Biaya operasional juga relatif rendah karena sistem dapat dijalankan pada pelanggan sederhana. Namun, pengembalian investasi jangka panjang rendah karena sistem tidak dapat diperluas dan akan memerlukan perancangan ulang besar ketika kebutuhan bisnis berkembang. Penghematan dari otomatisasi terbatas karena sistem hanya dapat menangani kueri yang sangat terstruktur dan sederhana, sehingga sebagian besar pekerjaan analisis masih memerlukan intervensi manual. Biaya peluang tinggi karena organisasi kehilangan peluang untuk mengoptimalkan pengambilan keputusan berbasis data secara cepat dan responsif.
<i>Customer</i> (Pengguna)	Pengalaman pengguna yang ditawarkan cukup terbatas. Pengguna harus menyesuaikan diri dengan format kueri yang kaku dan tidak fleksibel, yang tidak sesuai dengan cara berpikir dan berkomunikasi wajar mereka. Sistem tidak dapat memahami variasi pertanyaan atau konteks yang kompleks, sehingga sering kali pengguna harus mencoba beberapa kali dengan perumusan yang berbeda untuk mendapatkan hasil yang diinginkan. Hal ini menyebabkan frustrasi dan menurunkan tingkat kepuasan pengguna. Tingkat adopsi sistem cenderung rendah karena pengguna merasa sistem tidak memberikan nilai tambah yang signifikan dibandingkan dengan metode yang sudah ada.

Bersambung ke halaman berikutnya

Tabel III.4 – Analisis *Balanced Scorecard* untuk Aturan Sederhana (lanjutan)

Perspektif	Penilaian
<i>Internal Business Process</i>	Dari perspektif proses bisnis internal, solusi ini hanya menawarkan otomasi minimal untuk kueri yang sangat terstruktur dan berulang. Proses analisis manual masih tetap diperlukan untuk sebagian besar pertanyaan bisnis yang bersifat khusus atau kompleks, sehingga hambatan dalam alur analisis data tidak teratasi secara efektif. Waktu penyelesaian untuk mendapatkan wawasan bisnis tidak berkurang secara signifikan. Sistem juga tidak mendukung analisis lanjutan seperti penggalan data lebih dalam, perbandingan multi-dimensi, atau deteksi anomali, sehingga kemampuan organisasi untuk melakukan analisis mendalam tetap terbatas.
<i>Learning & Growth</i>	Solusi ini memberikan kesempatan yang sangat terbatas untuk pembelajaran dan pertumbuhan organisasi. Teknologi yang digunakan relatif sederhana dan tidak mendorong tim untuk mengembangkan keahlian baru dalam bidang kecerdasan buatan, pembelajaran mesin, atau analitik lanjutan. Organisasi tidak memperoleh keunggulan kompetitif teknologi dan tidak membangun kapabilitas yang dapat menjadi pembeda di pasar. Sistem juga tidak menyediakan landasan untuk inovasi masa depan atau integrasi dengan teknologi yang lebih canggih.

Solusi 2: *Chatbot* BI dengan Klasifikasi Maksud Berbasis Pola, Pembangkitan Bahasa Alami Berbasis Templat, dan Peramalan Deret Waktu

Tabel III.5 Analisis *Balanced Scorecard* untuk BI Komprehensif

Perspektif	Penilaian
<i>Financial</i>	<p>Solusi ini memerlukan investasi awal yang lebih tinggi, mencakup biaya untuk pengembangan sistem yang kompleks, pengadaan infrastruktur komputasi yang memadai (termasuk kemungkinan GPU untuk model peramalan), dan pelatihan model kecerdasan buatan. Namun, pengembalian investasi jangka panjang sangat tinggi karena sistem memberikan nilai tambah yang signifikan melalui beberapa jalur: (1) Pengurangan dramatis dalam waktu analisis dari jam atau hari menjadi detik, yang memungkinkan pengambilan keputusan yang lebih cepat dan responsif; (2) Penghematan biaya tenaga kerja karena analis dan tim TI tidak perlu menghabiskan waktu untuk kueri rutin dan dapat fokus pada tugas-tugas yang lebih strategis; (3) Peningkatan kualitas keputusan bisnis berdasarkan wawasan prediktif yang akurat, yang dapat menghasilkan peningkatan pendapatan atau penghematan biaya yang signifikan; (4) Skalabilitas yang baik, dengan biaya tambahan untuk menambahkan pengguna atau kueri baru relatif rendah setelah sistem terbangun. Sistem ini juga memberikan keunggulan kompetitif yang dapat berdampak pada posisi pasar dan profitabilitas jangka panjang organisasi.</p>

Bersambung ke halaman berikutnya

Tabel III.5 – Analisis *Balanced Scorecard* untuk BI Komprehensif (lanjutan)

Perspektif	Penilaian
<i>Customer</i> (Pengguna)	Solusi ini menawarkan pengalaman pengguna yang sangat unggul. Antarmuka percakapan yang wajar dan akrab memungkinkan pengguna untuk mengajukan pertanyaan dengan cara yang sama seperti mereka berbicara dengan rekan manusia, tanpa perlu mempelajari tata bahasa khusus atau terminologi teknis. Sistem dapat memahami berbagai variasi pertanyaan dan konteks, sehingga mengurangi frustrasi pengguna. Respons yang dihasilkan bersifat naratif, beragam, dan mudah dipahami, bukan hanya tabel angka yang kering. Kemampuan prediktif otomatis memberikan nilai tambah yang signifikan bagi pengguna, memungkinkan mereka untuk mengantisipasi tren dan membuat keputusan proaktif. Aksesibilitas berbagai platform (web, perangkat genggam, komputer pribadi) memastikan pengguna dapat mengakses wawasan setiap membutuhkannya. Tingkat kepuasan dan adopsi pengguna diperkirakan sangat tinggi karena sistem memberikan nilai nyata dan mudah digunakan.

Bersambung ke halaman berikutnya

Tabel III.5 – Analisis *Balanced Scorecard* untuk BI Komprehensif (lanjutan)

Perspektif	Penilaian
<i>Internal Business Process</i>	<p>Dari perspektif proses bisnis internal, solusi ini menawarkan transformasi yang fundamental. Otomasi maksimal untuk hampir semua jenis kueri analitik, dari yang sederhana hingga yang kompleks, menghilangkan hambatan yang selama ini menghalangi alur analisis data. Proses bisnis dapat menjadi lebih lincah dan responsif terhadap perubahan kondisi pasar atau internal. Waktu penyelesaian untuk mendapatkan wawasan berkurang drastis, memungkinkan organisasi untuk mengidentifikasi masalah, peluang, atau tren lebih cepat dan mengambil tindakan yang tepat. Kemampuan peramalan otomatis mendukung perencanaan yang lebih akurat untuk alokasi sumber daya, penetapan target, dan strategi bisnis. Sistem juga memfasilitasi budaya berbasis data dengan keputusan dibuat berdasarkan wawasan objektif bukan hanya intuisi.</p>

Bersambung ke halaman berikutnya

Tabel III.5 – Analisis *Balanced Scorecard* untuk BI Komprehensif (lanjutan)

Perspektif	Penilaian
<i>Learning & Growth</i>	Solusi ini memberikan kesempatan tertinggi untuk pembelajaran dan pertumbuhan organisasi. Implementasi teknologi kecerdasan buatan seperti pemrosesan bahasa alami, klasifikasi maksud berbasis pola, pembangkitan bahasa alami berbasis templat, dan peramalan deret waktu memungkinkan tim untuk mengembangkan keahlian baru yang sangat relevan di era digital. Organisasi membangun kapabilitas internal dalam bidang kecerdasan buatan dan pembelajaran mesin yang dapat menjadi fondasi untuk inovasi-inovasi masa depan, seperti ekspansi ke domain lain (<i>chatbot</i> layanan pelanggan, sistem rekomendasi, deteksi penipuan). Sistem menciptakan keunggulan kompetitif teknologi yang membedakan organisasi dari kompetitor. Platform yang dibangun bersifat dapat diperluas dan dapat menjadi basis untuk pengembangan solusi analitik yang lebih canggih di masa depan. Budaya inovasi dan adopsi teknologi terdorong, memposisikan organisasi sebagai pemimpin dalam transformasi digital di industrinya.

Solusi 3: Dasbor BI Interaktif Tradisional dengan Peningkatan

Tabel III.6 Analisis *Balanced Scorecard* untuk BI Tradisional

Perspektif	Penilaian
<i>Financial</i>	Solusi ini memerlukan investasi yang sedang, terutama untuk meningkatkan antarmuka pengguna dan menambahkan fitur interaktivitas pada dasbor yang sudah ada. Biaya lisensi untuk alat BI komersial dapat cukup signifikan, terutama jika organisasi menggunakan platform seperti Tableau, Power BI, atau Qlik. Pengembalian investasi berada di tingkat menengah karena peningkatan efisiensi terbatas pada perbaikan antarmuka pengguna dan kemudahan akses, tanpa transformasi fundamental dalam cara pengguna berinteraksi dengan data. Penghematan biaya tenaga kerja terbatas karena pengguna masih memerlukan keahlian teknis untuk memanfaatkan fitur analitik layanan mandiri secara efektif. Tidak ada kemampuan prediktif otomatis yang dapat memberikan nilai tambah signifikan untuk perencanaan strategis.
<i>Customer</i> (Pengguna)	Pengalaman pengguna mengalami peningkatan dibandingkan dengan sistem lama, terutama dalam hal visualisasi data dan kemudahan navigasi. Dasbor yang lebih interaktif dengan fitur penggalian data lebih dalam dan penyaringan lanjutan memberikan fleksibilitas lebih besar bagi pengguna untuk mengeksplorasi data. Aplikasi perangkat genggam memungkinkan akses data dari perangkat bergerak, meningkatkan kemudahan akses. Namun, solusi ini masih tidak mendukung kueri bahasa wajar atau antarmuka percakapan, sehingga pengguna nonteknis tetap menghadapi kurva pembelajaran yang cukup tinggi. Sistem tidak dapat menjawab pertanyaan khusus yang kompleks tanpa konfigurasi tambahan dari tim TI. Kepuasan pengguna berada di tingkat menengah karena peningkatan yang diberikan bersifat bertahap, bukan transformatif.

Bersambung ke halaman berikutnya

Tabel III.6 – Analisis *Balanced Scorecard* untuk BI Tradisional (lanjutan)

Perspektif	Penilaian
<i>Internal Business Process</i>	Dari perspektif proses bisnis, solusi ini memberikan peningkatan yang sedang hingga tinggi dalam efisiensi operasional. Pengguna yang memiliki keahlian teknis dapat melakukan analisis sendiri tanpa selalu bergantung pada tim TI, mengurangi beban kerja analis. Fitur analitik layanan mandiri memungkinkan eksplorasi data yang lebih fleksibel untuk pengguna yang terlatih. Namun, untuk pengguna nonteknis atau pertanyaan yang sangat spesifik dan kompleks, sistem masih memerlukan intervensi manual dari tim analitik. Tidak ada dukungan untuk analisis prediktif otomatis, sehingga organisasi tidak dapat memanfaatkan kemampuan peramalan untuk perencanaan strategis yang lebih baik. Proses tetap tidak dapat diperluas untuk kebutuhan bisnis yang sangat dinamis dan peka waktu.
<i>Learning & Growth</i>	Solusi ini memberikan kesempatan sedang untuk pembelajaran dan pertumbuhan. Organisasi dapat meningkatkan literasi data tim melalui pelatihan penggunaan alat BI dan fitur analitik layanan mandiri. Penguasaan alat BI populer seperti Tableau atau Power BI dapat menjadi keahlian yang berharga bagi karyawan. Namun, solusi ini tidak mendorong inovasi teknologi yang signifikan atau pengembangan kapabilitas dalam bidang kecerdasan buatan dan pembelajaran mesin. Organisasi tidak membangun pembeda teknologi yang kuat dan tetap bergantung pada vendor eksternal untuk sebagian besar fitur dan inovasi. Platform ini juga terbatas dalam hal kemampuan perluasan untuk pengembangan solusi khusus atau integrasi dengan teknologi yang lebih canggih di masa depan.

Solusi 4: Chatbot BI Berbasis Model Bahasa Besar (Large Language Model)

Tabel III.7 Analisis *Balanced Scorecard* untuk LLM

Perspektif	Penilaian
<i>Financial</i>	<p>Solusi ini memerlukan investasi awal yang sangat tinggi, mencakup biaya lisensi atau akses API untuk model bahasa besar komersial (seperti GPT-4 atau Claude), pengadaan infrastruktur komputasi dengan GPU berkinerja tinggi untuk inferensi cepat, dan biaya pengembangan untuk integrasi sistem yang kompleks. Biaya operasional juga sangat signifikan karena setiap kueri memerlukan sumber daya komputasi yang besar, terutama untuk model dengan miliaran parameter. Untuk model berbasis API, biaya per token dapat terakumulasi dengan cepat pada skala organisasi besar. Pengembalian investasi berada pada tingkat sedang hingga tinggi dalam jangka panjang, bergantung pada seberapa baik organisasi dapat mengelola risiko halusinasi dan memastikan akurasi respons. Nilai tambah utama berasal dari kemampuan pemahaman bahasa yang sangat canggih dan fleksibilitas dalam menangani kueri kompleks yang tidak terstruktur. Namun, ketidakpastian dalam kontrol keluaran dan risiko kesalahan dapat membatasi pengembalian investasi jika tidak dikelola dengan baik. Biaya peluang juga perlu dipertimbangkan jika organisasi mengalami insiden kebocoran data atau respons yang tidak akurat yang berdampak pada keputusan bisnis.</p>

Bersambung ke halaman berikutnya

Tabel III.7 – Analisis *Balanced Scorecard* untuk LLM (lanjutan)

Perspektif	Penilaian
<i>Customer</i> (Pengguna)	<p>Solusi ini menawarkan pengalaman pengguna yang sangat unggul dalam hal kewajaran interaksi dan fleksibilitas komunikasi. Pengguna dapat mengajukan pertanyaan dengan cara yang sangat wajar, termasuk kueri yang kompleks, ambigu, atau memerlukan pemahaman konteks mendalam yang sulit ditangani sistem konvensional. Model bahasa besar dapat memahami nuansa, konteks percakapan multiturn, dan bahkan inferensi implisit dari pertanyaan pengguna. Respons yang dihasilkan sangat naratif, koheren, dan menyerupai komunikasi manusia, yang meningkatkan kepuasan pengguna secara signifikan. Namun, terdapat risiko bahwa sistem dapat menghasilkan respons yang terdengar meyakinkan tetapi faktanya tidak akurat (halusinasi), yang dapat menyesatkan pengguna dan menurunkan kepercayaan terhadap sistem. Pengguna nonteknis mungkin tidak dapat membedakan respons yang akurat dari yang dihasilkan secara spekulatif. Untuk mempertahankan kepuasan tinggi, sistem memerlukan mekanisme validasi dan verifikasi yang ketat, serta transparansi mengenai tingkat keyakinan respons.</p>

Bersambung ke halaman berikutnya

Tabel III.7 – Analisis *Balanced Scorecard* untuk LLM (lanjutan)

Perspektif	Penilaian
<i>Internal Business Process</i>	<p>Dari perspektif proses bisnis internal, solusi ini memberikan fleksibilitas tertinggi dalam menangani berbagai jenis kueri analitik, termasuk yang sangat kompleks, tidak terstruktur, atau belum pernah ditemui sebelumnya. Kemampuan model untuk memahami konteks yang dalam dan melakukan inferensi dapat mengurangi kebutuhan untuk konfigurasi manual atau pemrograman khusus untuk kasus penggunaan baru. Sistem dapat beradaptasi dengan perubahan kebutuhan bisnis dengan lebih organik. Namun, ketidakdeterministikan keluaran model bahasa besar menimbulkan tantangan dalam proses bisnis yang memerlukan konsistensi dan akurasi tinggi. Risiko halusinasi dapat menyebabkan keputusan bisnis yang salah jika tidak ada mekanisme validasi yang memadai. Latensi respons yang lebih tinggi dibandingkan solusi berbasis pola juga dapat menjadi hambatan untuk proses bisnis yang sangat sensitif terhadap waktu. Integrasi dengan sistem keamanan dan tata kelola data memerlukan perhatian khusus untuk memastikan bahwa model tidak secara tidak sengaja mengekspos informasi sensitif atau melanggar kebijakan privasi.</p>

Bersambung ke halaman berikutnya

Tabel III.7 – Analisis *Balanced Scorecard* untuk LLM (lanjutan)

Perspektif	Penilaian
<i>Learning & Growth</i>	Solusi ini memberikan kesempatan sangat tinggi untuk pembelajaran dan pertumbuhan organisasi dalam bidang teknologi mutakhir. Implementasi model bahasa besar memungkinkan tim untuk mengembangkan keahlian dalam pembelajaran mendalam, arsitektur Transformer, penyetelan-halus model, dan teknik prompt engineering yang merupakan kompetensi sangat relevan di era kecerdasan buatan generatif. Organisasi dapat memposisikan diri sebagai pelopor dalam adopsi teknologi kecerdasan buatan paling canggih. Namun, ketergantungan pada model dan infrastruktur dari penyedia eksternal (seperti OpenAI atau Anthropic) dapat membatasi kontrol organisasi terhadap teknologi inti dan menimbulkan risiko vendor lock-in. Jika menggunakan model sumber terbuka, organisasi memerlukan investasi signifikan dalam keahlian spesialisasi untuk pelatihan, penyetelan, dan pemeliharaan model. Platform ini membuka peluang untuk inovasi masa depan seperti asisten virtual multidomain, sistem rekomendasi kontekstual, atau analitik prediktif yang sangat canggih, tetapi memerlukan komitmen jangka panjang dalam pengembangan kapabilitas dan manajemen risiko.

Matriks Perbandingan *Balanced Scorecard*

Untuk memberikan gambaran yang lebih komprehensif, berikut adalah matriks perbandingan ketiga solusi berdasarkan keempat perspektif *Balanced Scorecard*, beserta karakteristik pengguna yang paling sesuai untuk masing-masing solusi:

Tabel III.8 Perbandingan *Balanced Scorecard* Alternatif Solusi

Solusi	<i>Financial</i>	<i>Customer</i>	<i>Process</i>	<i>L&G</i>	Digunakan oleh
<i>Chatbot</i> Berbasis Aturan Sederhana	Rendah- Sedang	Rendah- Sedang	Rendah	Rendah	Organisasi kecil dengan kebu- tuhan analitik sangat terbatas dan kueri yang sangat terstruktur
<i>Chatbot</i> BI dengan Klasifikasi Maksud, Pem- bangkitan Bahasa Alami, dan Peramal- an Deret Waktu	Sedang- Tinggi (ROI Tinggi)	Sangat Tinggi	Sangat Tinggi	Sangat Tinggi	Organisasi me- nengah hingga besar dengan volume data tinggi, kebutuhan analitik kom- pleks, dan fokus pada transformasi digital berbasis AI
Dasbor BI Interaktif Tradisional dengan Pe- ningkatan	Sedang	Sedang	Sedang- Tinggi	Sedang	Organisasi yang sudah memiliki investasi besar dalam alat BI tra- disional dan ingin peningkatan bertahap

Bersambung ke halaman berikutnya

Tabel III.8 – Perbandingan *Balanced Scorecard* Alternatif Solusi (lanjutan)

Solusi	<i>Financial</i>	<i>Customer</i>	<i>Process</i>	<i>L&G</i>	Digunakan oleh
Chatbot BI Berbasis Model Bahasa Besar (LLM)	Sedang-Tinggi (Biaya Tinggi)	Sangat Tinggi (dengan Risiko)	Tinggi (dengan Tantangan)	Sangat Tinggi	Organisasi besar dengan anggaran signifikan, infrastruktur AI matang, kebutuhan pemahaman bahasa sangat kompleks, dan kapabilitas manajemen risiko AI yang kuat

Kesimpulan Pemilihan Solusi

Berdasarkan analisis komprehensif menggunakan metode *Balanced Scorecard* pada keempat perspektif (*Financial*, *Customer*, *Internal Business Process*, dan *Learning & Growth*), **Solusi 2: Chatbot BI dengan Klasifikasi Maksud Berbasis Pola, Pembangkitan Bahasa Alami Berbasis Templat, dan Peramalan Deret Waktu** adalah solusi yang paling optimal dan direkomendasikan untuk diimplementasikan. Perlu dicatat bahwa Solusi 4 (Chatbot BI Berbasis Model Bahasa Besar) menawarkan kemampuan pemahaman bahasa alami yang paling canggih dan pengalaman pengguna yang sangat unggul. Namun, solusi ini memiliki beberapa kelemahan kritis yang membuatnya kurang optimal untuk konteks penelitian ini: (1) Biaya implementasi dan operasional yang sangat tinggi, yang dapat tidak proporsional dengan kebutuhan organisasi menengah; (2) Risiko halusinasi dan respons tidak akurat yang dapat berdampak serius pada keputusan bisnis berbasis data; (3) Kesulitan dalam mengontrol keluaran secara deterministik, yang penting untuk sistem BI yang memerlukan konsistensi dan akurabilitas; (4) Potensi kebocoran data sensitif melalui model yang tidak sepenuhnya dikontrol organisasi; dan (5) Ketergantungan pada penyedia model eksternal yang menimbulkan risiko keberlanjutan jangka panjang. Meskipun teknologi ini sangat menjanjikan untuk masa depan, untuk implementasi praktis dalam konteks organisasi yang memerlukan keseimbangan antara inovasi dan stabilitas, Solusi 2 tetap menjadi pilihan yang paling optimal. Alasan utama pemilihan Solusi 2 adalah sebagai berikut:

1. **Keselarasan dengan Kebutuhan:** Solusi 2 secara langsung mengatasi semua

masalah pengguna yang telah diidentifikasi, termasuk keterlambatan respons, kesulitan akses analisis khusus, ketiadaan kemampuan prediksi otomatis, dan akses yang tidak ramah pengguna nonteknis. Solusi ini memenuhi seluruh kebutuhan fungsional dan nonfungsional yang telah disusun secara komprehensif.

2. **Skalabilitas dan Adaptabilitas:** Sistem yang diusulkan dirancang dengan arsitektur modular dan dapat diperluas yang dapat diadaptasi terhadap perubahan kebutuhan bisnis di masa depan tanpa memerlukan perancangan ulang besar. Kemampuan untuk menambahkan maksud baru, aturan baru, atau model peramalan baru secara mudah memberikan fleksibilitas jangka panjang.
3. **Tahan Masa Depan dan Platform untuk Inovasi:** Platform yang dibangun membuka pintu untuk inovasi-inovasi masa depan seperti ekspansi ke domain lain (*chatbot* layanan pelanggan, sistem rekomendasi, deteksi penipuan), integrasi dengan sistem eksternal (ERP, CRM), atau pengembangan kemampuan analitik lanjutan lainnya. Investasi dalam platform ini tidak hanya memberikan nilai untuk kebutuhan saat ini, tetapi juga menjadi fondasi untuk transformasi digital organisasi secara keseluruhan.
4. **Pengembalian Investasi dan Justifikasi Finansial:** Meskipun investasi awal lebih tinggi dibandingkan alternatif lain, pengembalian investasi jangka panjang sangat menarik karena penghematan biaya tenaga kerja yang signifikan (analisis dapat fokus pada tugas strategis), pengambilan keputusan yang lebih cepat dan akurat (yang dapat berdampak langsung pada pendapatan atau penghematan biaya), dan peningkatan daya saing organisasi. Berdasarkan studi dari berbagai industri, implementasi sistem analitik prediktif berbasis kecerdasan buatan dapat meningkatkan efisiensi operasional hingga 30% dan mempercepat pengambilan keputusan hingga 85% (Shah, 2025).
5. **Keunggulan Kompetitif dan Diferensiasi:** Implementasi *chatbot* BI cangih dengan kemampuan kecerdasan buatan dan pembelajaran mesin memberikan keunggulan kompetitif bagi organisasi, meningkatkan kelincahan dan kecepatan pengambilan keputusan dibandingkan kompetitor yang masih menggunakan sistem konvensional. Organisasi yang mampu menganalisis data dan merespons tren pasar lebih cepat memiliki posisi yang lebih kuat dalam persaingan bisnis.
6. **Adopsi Pengguna dan Kepuasan:** Antarmuka percakapan berbasis bahasa wajar memiliki tingkat adopsi yang jauh lebih tinggi dibandingkan dashboard tradisional karena lebih intuitif dan tidak memerlukan pelatihan ekstensif. Statistik industri menunjukkan bahwa *chatbot* dapat menangani hingga 80%

pertanyaan standar secara otomatis (Fullview, 2025), yang secara signifikan mengurangi beban kerja tim dan meningkatkan kepuasan pengguna.

7. **Keamanan dan Kontrol Data:** Pendekatan pembangkitan bahasa alami berbasis templat memberikan kontrol penuh atas respons yang dihasilkan sistem, memastikan bahwa tidak ada kebocoran data sensitif atau informasi yang tidak seharusnya diungkapkan. Hal ini sangat penting untuk sistem yang beroperasi dengan data internal perusahaan yang bersifat rahasia.
8. **Pembelajaran dan Pertumbuhan Organisasi:** Implementasi solusi ini memungkinkan organisasi untuk membangun kapabilitas internal dalam bidang kecerdasan buatan, pembelajaran mesin, dan analitik lanjutan, yang merupakan keahlian kritis di era digital. Tim akan mengembangkan pengetahuan dan pengalaman yang berharga yang dapat diterapkan pada proyek-proyek inovasi lainnya di masa depan.

Dengan pemilihan Solusi 2, organisasi akan dapat mencapai transformasi digital yang signifikan dalam domain *Business Intelligence* dan analitik, meningkatkan efisiensi operasional, kecepatan pengambilan keputusan, dan memposisikan diri sebagai pemimpin dalam pemanfaatan teknologi kecerdasan buatan untuk keunggulan kompetitif di pasar yang semakin berbasis data.

Solusi yang dipilih juga sejalan dengan tren global dengan pasar kecerdasan buatan percakapan diproyeksikan mencapai USD 49,7 miliar pada tahun 2025 (Qaltivate, 2025), pasar pemrosesan bahasa alami tumbuh dengan CAGR 28,6% mencapai USD 438,08 miliar pada tahun 2034 (Precedence Research, 2025), dan pasar peramalan deret waktu diperkirakan mencapai USD 36,9 miliar pada tahun 2032 (WiseGuy Reports, 2024). Pertumbuhan pasar yang eksponensial ini menunjukkan bahwa investasi dalam teknologi-teknologi ini bukan hanya relevan untuk saat ini, tetapi juga strategis untuk jangka panjang.

DAFTAR PUSTAKA

- Alharbi, Nada. 2025. "Prediction of Stock Market Using LSTM, ARIMA, and Hybrid ARIMA-LSTM Models: A Comparative Analysis". *Applied Mathematics & Information Sciences* 19 (1): 45–58. <https://doi.org/10.18576/amis/190105>.
- Amazon Web Services. 2024. "Securing Generative AI: Data, Compliance, and Privacy Considerations". Diakses pada November 4, 2025. <https://aws.amazon.com/blogs/security/securing-generative-ai-data-compliance-and-privacy-considerations/>.
- Al-Amri, Rashid, R. K. Murugesan, dkk. 2025. "Enhancing Customer Churn Analysis by Using Real-time Continual Learning". *International Journal of Advanced Computer Science and Applications* 16 (5): 1–12.
- AtScale. 2025. "Why Semantic Layers? A Key Technical Architecture". Diakses pada November 4, 2025. <https://www.atscale.com/blog/semantic-layers-key-technical-architecture/>.
- Bamurange, Daniel. 2025. "Designing a Hybrid AI Chatbot Framework for Student Support: Integrating NLP and Human Oversight in African Universities". *Edinburgh Journals - Journal of Information Technology* 5 (4): 41–52. <https://doi.org/10.70619/vol5iss4pp41-52>.
- Business Wire. 2020. "The World's AI-Based Chatbot Market, 2020-2025: Conversational AI Will Improve Telemarketing Efficiency". Diakses pada November 4, 2025. <https://www.businesswire.com/news/home/20200325005452/en/>.
- Cyara. 2024. "What's the Optimum Confidence Threshold for My Chatbot". Diakses pada November 4, 2025. <https://cyara.com/blog/optimum-chatbot-confidence-threshold/>.

- Databricks. 2025. “The Role of Semantic Layers in Modern Data Analytics”. Diakses pada November 4, 2025. <https://www.databricks.com/glossary/semantic-layer>.
- Deemter, Kees van, Emiel Krahmer, dan Mariët Theune. 2005. “Real versus Template-based Natural Language Generation: A False Opposition?” *Computational Linguistics* 31 (1): 15–24. <https://doi.org/10.1162/0891201053630291>.
- Dubey, Ashish Kumar, Abhishek Kumar, dan Sandeep Narang. 2021. “Study and Analysis of SARIMA and LSTM in Forecasting Time Series Data”. *Sustainable Energy Technologies and Assessments* 47:101474. <https://doi.org/10.1016/j.seta.2021.101474>.
- East Ventures. July 2025. *AI-first: Decoding Southeast Asia Trends*. White Paper. Diakses: 4 November 2025. Jakarta, Indonesia: East Ventures. <https://east.vc/press-release-3/east-ventures-launched-a-white-paper-focusing-on-ai>.
- European Data Protection Board. 2025. “AI Privacy Risks & Mitigations: Large Language Models (LLMs)”. Diakses pada November 4, 2025. <https://www.edpb.europa.eu/>.
- Evidently AI. 2025. “Accuracy, Precision, and Recall in Multi-Class Classification”. Diakses pada December 2, 2025. <https://www.evidentlyai.com/classification-metrics/multi-class-metrics>.
- Feyisetan, Oluwaseyi, Borja Balle, Thomas Drake, dan Tom Diethe. 2022. “Privacy and Utility-Preserving Textual Analysis via Calibrated Multivariate Perturbations”. *Proceedings of the 13th ACM International Conference on Web Search and Data Mining*, 178–186. <https://doi.org/10.1145/3336191.3371856>.
- Fujitsu - PostgreSQL Fastware. 2024. “Generative AI and Data: Protecting Organizational Data”. Diakses pada November 4, 2025. <https://www.postgresql.fastware.com/blog/generative-ai-and-data>.
- Fullview. 2025. “80+ AI Customer Service Statistics & Trends in 2025”. Diakses pada November 4, 2025. <https://www.fullview.io/blog/ai-customer-service-stats>.

- Gehrmann, Sebastian, Tosin Adewumi, Karmanya Aggarwal, Pawan Sasanka Amanamanchi, Anuoluwapo Aremu, Antoine Bosselut, dkk. 2022. “NLG-Metricverse: An End-to-End Library for Evaluating Natural Language Generation”. *Proceedings of the 29th International Conference on Computational Linguistics*, 3465–3475.
- Gehweiler, Christian. 2024. “Classification of Intent in Moderating Online Discussions”. *Expert Systems with Applications* 239:122234.
- Gehweiler, Christian, dkk. 2025. “Intent Detection for Task-Oriented Conversational Agents: A Survey”. *Expert Systems* 42 (3).
- Genesys. 2023. “Set Bot Confidence Thresholds”. Diakses pada November 4, 2025. <https://www.genesys.com/blog/post/set-bot-confidence-thresholds>.
- Grandini, Margherita, Enrico Bagli, dan Giorgio Visani. 2020. “Metrics for Multi-Class Classification: An Overview”. *arXiv preprint arXiv:2008.05756*, arXiv: 2008.05756 [cs.LG].
- Huang, Xufei. 2021. “Chatbot: Design, Architecture, and Applications”. PhD thesis, University of Pennsylvania. <https://www.cis.upenn.edu/wp-content/uploads/2021/10/Xufei-Huang-thesis.pdf>.
- Intimedia. Desember 2024. “Indonesia’s Digital Transformation in 2024: How AI is Turning Dreams into Reality”. Diakses: 4 November 2025, *Intimedia News* (). <https://intimedia.id/read/indonesias-digital-transformation-in-2024-how-ai-is-turning-dreams-into-reality>.
- Iqbal, Muhammad, Rashid Ahmad, dan Saif Ur Rehman Khan. 2021. “Combined Support Vector Machine and Pattern Matching for Arabic Question Classification”. *International Journal of Pattern Recognition and Artificial Intelligence* 35 (8): 287–299. https://doi.org/10.1007/978-3-319-99007-1_27.
- IT Convergence. 2024. “Generative AI Data Security: Key Considerations”. Diakses pada November 4, 2025. <https://www.itconvergence.com/blog/data-security-considerations-for-generative-ai/>.
- Kale, Mihir, dan Abhinav Rastogi. 2020. “Template Guided Text Generation for Task-Oriented Dialogue”. Dalam *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 6505–6520. Association for Computational Linguistics. <https://doi.org/10.18653/v1/2020.emnlp-main.527>.

- Kapoor, Ankit, dan Sujala Shetty. 2023. "Implementing Natural Language Generation Through Industry-Specific Chatbots". *Data Science and Intelligent Computing Techniques*, 55–67. <https://doi.org/10.56155/978-81-955020-2-8-6>.
- Khaddam, Ibrahim, dan Mohammad Alzghoul. 2025. "Artificial Intelligence-Driven Business Intelligence: Systematic Literature Review". *International Journal of Energy Economics and Policy* 15 (1): 45–54. <https://doi.org/10.32479/ijeep.19820>.
- Khosla, Sopan, Chenwei Xu, Shikhar Toshniwal, Dinesh Manocha, Markus Dreyer, Lambert Mathias, Jing Gao, dan Asli Celikyilmaz. 2022. "Evaluating the Practical Utility of Confidence-score based Techniques for Unsupervised Open-World Classification". Dalam *Proceedings of the Third Workshop on Insights from Negative Results in NLP*, 20–28. <https://doi.org/10.18653/v1/2022.insights-1.3>.
- Kuligowska, Karolina, dan Zdzislaw Kowalczyk. 2024. "Enhancing Chatbot Intent Classification Using Active Learning Pipeline". *Journal of Applied Economic Sciences* 19 (2): 290–300.
- Kumar, Arnav, dan Ebad Zafar. 2024. "Predict Customer Churn with Python and Machine Learning". *SSRN Electronic Journal*, <https://doi.org/10.2139/ssrn.5085192>.
- Labs, dbt. 2024. "Understanding Semantic Layer Architecture". Diakses pada November 4, 2025. <https://www.getdbt.com/blog/semantic-layer-architecture>.
- Li, Xiaoming, Yutao Zhang, Yang Liu, dan Wei Wang. 2024. "Customer Churn Prediction Model Based on Hybrid Neural Networks". *Scientific Reports* 14:79603. <https://doi.org/10.1038/s41598-024-79603-9>.
- Luo, Yuanshuai, dan Guiran Liu. 2024. "Data Security Strategies in Natural Language Processing Applications". *arXiv preprint arXiv:2410.08553*, arXiv: 2410.08553.
- Mena, C. G., A. De Caigny, K. Coussement, K. W. De Bock, dan S. Lessmann. 2023. "Exploiting time-varying RFM measures for customer churn prediction with deep neural networks". Diakses: Tanggal Akses Anda, *Annals of Operations Research* 339 (1-2): 765–787. URL_ATAU_DOI_DISINI.

- Merugu, Srujana, dkk. 2024. "Intent Detection in the Age of LLMs". Dalam *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing: Industry Track*, 1–12.
- Mireshghallah, Fatemehsadat, Kartik Goyal, Shrimai Upadhyay, dan Taylor Berg-Kirkpatrick. 2022. "Differentially Private Methods in Natural Language Processing". *Proceedings of the Fourth Workshop on Privacy in Natural Language Processing*, 1–11. <https://doi.org/10.18653/v1/2022.privatenlp-1.1>.
- Mulyatun, Dedi Kurniawan, dan Agus Hudaya. 2021. "Pendekatan Natural Language Processing pada Aplikasi Chatbot sebagai Alat Bantu Customer Service". *Jurnal Teknologi Sistem Informasi dan Aplikasi* 4 (1): 55–62. <https://doi.org/10.32493/jtsi.v4i1.7751>.
- Nabiilah, Sarah, Adiwijaya, dan Widi Maharani. 2025. "Comparative Analysis of SVM and IndoBERT for Intent Classification in E-LEMBUR Chatbot". *Journal of Software and Computer Engineering* 6 (3): 258–267. <https://doi.org/10.61628/jsce.v6i3.2058>.
- Ni'mah, Iftitahu, Meng Fang, Vlado Menkovski, dan Mykola Pechenizkiy. 2023. "NLG Evaluation Metrics Beyond Correlation Analysis: An Empirical Metric Preference Checklist". *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics*, 1240–1266. <https://doi.org/10.18653/v1/2023.acl-long.69>.
- Ouaddi, Chaimae, Fatima Benabbou, dan Nawal Sael. 2025. "A Comparative Analysis and Performance Evaluation of Deep Learning Techniques for Intent Detection in Conversational AI Chatbots". *Journal of Big Data* 12 (1): 1–25. <https://doi.org/10.1186/s40537-024-00996-6>.
- Ozerova, Daria. 2022. "Intent Detection Module for a Conversational Assistant". Master's thesis, Czech Technical University in Prague.
- Pecan AI. 2024. "Best Machine Learning Models for Predicting Customer Churn". Diakses pada November 4, 2025. <https://www.pecan.ai/blog/best-ml-models-for-predicting-customer-churn/>.
- Precedence Research. 2025. "Natural Language Processing Market Size to Hit USD 438.08 Bn by 2034". Diakses pada November 4, 2025. <https://www.precedenceresearch.com/natural-language-processing-market>.

- Qaltivate. 2025. "Conversational AI Market Summary: 2025 Edition". Diakses pada November 4, 2025. <https://qaltivate.com/blog/conversational-ai-market/>.
- Rahmat, Budi, dan Agus Setiawan. 2024. "Prediction Model Using Machine Learning: Analysis of Determinants of Customer Churn at PT XYZ". *Jurnal Info Sains: Informatika dan Sains* 14 (04): 745–758.
- Rebuffel, Clément, Laure Soulier, Geoffrey Scoutheeten, dan Patrick Gallinari. 2020. "A Hierarchical Model for Data-to-Text Generation". *arXiv preprint arXiv:2004.15006*, arXiv: 2004.15006.
- Santosa, Rahmad, Adetiya Bagus Nusantara, dan Syaiful Imron. 2025. "Comparative Analysis of SVM and IndoBERT for Intent Classification in Indonesian Overtime Chatbots". *JSCE* 6 (3): 258–270. <https://doi.org/10.61628/jsce.v6i3.2058>.
- Shah, Nisarg. 2025. "Hybrid Analytics Architecture: Integrating Traditional BI with AI-Driven Insights". *World Journal of Advanced Engineering Technology and Sciences* 14 (02): 256–271. <https://doi.org/10.30574/wjaets.2025.14.2.0037>.
- Sharma, Rajesh, Amit Gupta, dan Pradeep Singh. 2025. "Customer Churn Prediction: A Machine Learning Approach with Data Balancing for Telecom Industry". *International Journal of Computing* 24 (1): 1–12. <https://doi.org/10.47839/ijc.24.1.3873>.
- Siami-Namini, Sima, Neda Tavakoli, dan Akbar Siami Namin. 2018. "A Comparison of ARIMA and LSTM in Forecasting Time Series". Dalam *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*, 1394–1401. IEEE. <https://doi.org/10.1109/ICMLA.2018.00227>.
- Sirisha, U. M., M. C. Belavagi, dan G. Attigeri. 2022. "Profit Prediction Using ARIMA, SARIMA and LSTM Models in Time Series Forecasting: A Comparison". *International Journal of Current Research and Review (IJCRR)* 14 (24): 32–42. <https://doi.org/10.31782/IJCRR.2022.142410>.
- Sorour, Mohamed K. 2020. "The Role of Business Intelligence and Analytics in Higher Education Decision Making". Dalam *2020 IEEE Global Engineering Education Conference (EDUCON)*, 1849–1856. IEEE. <https://doi.org/10.1109/EDUCON45650.2020.9194157>.

- Spot Intelligence. 2023. "Building a Dataset for Intent Classification in NLP: Evaluation and Model Performance". Diakses pada November 4, 2025. <https://spotintelligence.com/2023/11/03/intent-classification-nlp/>.
- Straits Research. 2024. "Business Intelligence Market Size & Outlook, 2025–2033". Diakses pada November 4, 2025. <https://straitsresearch.com/report/business-intelligence-market>.
- Sunendar, Agus, Sunu Wibirama, dan Noor Akhmad Setiawan. 2025. "Comparison of ARIMA, LSTM, and GRU Models for Sales Prediction in Indonesian Banking Stock Market". *PILAR: Physics, Informatics, and Applied Research* 3 (1): 67–78. <https://doi.org/10.25273/pilar.v3i1.6412>.
- Teneo Developers. 2022. "Intent Classification: Confidence Threshold". Diakses pada November 4, 2025. <https://developers.teneo.ai/documentation/7.5.0/reference/conceptual-overviews/intent-classification>.
- Velotix. 2025. "Data Security in Generative AI Environments: 5 Steps". Diakses pada November 4, 2025. <https://www.velotix.ai/resources/blog/data-security-in-generative-ai-environments/>.
- WiseGuy Reports. 2024. "Time Series Forecasting Market". Diakses pada November 4, 2025. <https://www.wiseguyreports.com/reports/time-series-forecasting-market>.
- XenonStack. 2023. "Chatbots for Business: Best Practices and Top Applications". Diakses pada November 4, 2025. <https://www.xenonstack.com/insights/chatbots-applications>.
- Zhang, Yutao, Yang Liu, Meng Zhang, dan Xiaoming Li. 2022. "Disentangling Confidence Score Distribution for Out-of-Domain Intent Detection". *arXiv preprint arXiv:2210.08830*, arXiv: 2210.08830.