

**JUDUL TESIS UNTUK PROGRAM MAGISTER (S2)
DATA SCIENCE DAN KECERDASAN BUATAN**

TESIS

**NAMA MAHASISWA
NIM 000000000**



**PROGRAM STUDI S2 DATA SCIENCE DAN KECERDASAN BUATAN
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
UNIVERSITAS SUMATERA UTARA
MEDAN
2026**

**JUDUL TESIS UNTUK PROGRAM MAGISTER (S2)
DATA SCIENCE DAN KECERDASAN BUATAN**

TESIS

Diajukan untuk melengkapi tugas dan memenuhi syarat memperoleh ijazah
Magister Data Science dan Kecerdasan Buatan

NAMA MAHASISWA

NIM 000000000



**PROGRAM STUDI S2 DATA SCIENCE DAN KECERDASAN BUATAN
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
UNIVERSITAS SUMATERA UTARA
MEDAN
2026**

PERSETUJUAN

Judul : Judul Tesis untuk Program Magister (S2) Data Science dan Kecerdasan Buatan
Kategori : Tesis
Nama : Nama Mahasiswa
Nomor Induk Mahasiswa : 000000000
Program Studi : S2 Data Science dan Kecerdasan Buatan
Fakultas : Ilmu Komputer dan Teknologi Informasi
Universitas Sumatera Utara
Komisi Pembimbing :

Pembimbing 2

Pembimbing 1

Nama Lengkap dan Gelar Akademik

Nama Lengkap dan Gelar Akademik

Diketahui/disetujui oleh
Program Studi S2 Data Science dan Kecerdasan Buatan
Ketua,

Nama Lengkap dan Gelar Akademik
Nomor Induk Pegawai

PERNYATAAN

Judul Tesis untuk Program Magister (S2) Data Science dan Kecerdasan Buatan

TESIS

Saya mengakui bahwa tesis ini adalah hasil karya saya sendiri, kecuali beberapa kutipan dan ringkasan yang masing-masing telah disebutkan sumbernya.

Medan, Hari Bulan Tahun

Nama Mahasiswa

Nomor Induk Mahasiswa

**PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS**

Sebagai sivitas akademika Universitas Sumatera Utara, saya yang bertanda tangan di bawah ini:

Nama : Nama Mahasiswa
NIM : 0000000000
Program Studi : S2 Data Science dan Kecerdasan Buatan
Jenis Karya Ilmiah : Tesis

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Sumatera Utara Hak Bebas Royalti Non-Eksklusif (*Non-Exclusive Royalty Free Right*) atas tesis saya yang berjudul:

Judul Tesis untuk Program Magister (S2) Data Science dan Kecerdasan Buatan

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Non-Eksklusif ini, Universitas Sumatera Utara berhak menyimpan, mengalih media, memformat, mengelola dalam bentuk *database*, merawat dan mempublikasikan tesis saya tanpa meminta izin dari saya selama tetap mencantumkan nama saya sebagai penulis dan sebagai pemegang dan/atau sebagai pemilik hak cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Medan, hh-bulan-tahun

Nama Jelas
NIM

Telah diuji pada

Tanggal: hh bulan tahun _____

PANITIA PENGUJI TESIS

Ketua : Nama Lengkap dan Gelar Akademik

Anggota : 1. Nama Lengkap dan Gelar Akademik
2. Nama Lengkap dan Gelar Akademik
3. Nama Lengkap dan Gelar Akademik
4. Nama Lengkap dan Gelar Akademik

RIWAYAT HIDUP

DATA PRIBADI

Nama Lengkap (berikut gelar) :
Tempat dan Tanggal Lahir :
Alamat Rumah :
Telepon/Faks/HP :
E-mail :
Instansi Tempat Bekerja :
Alamat Kantor :

DATA PENDIDIKAN

SD	:	TAMAT: tahun
SLTP	:	TAMAT: tahun
SLTA	:	TAMAT: tahun
S1	:	TAMAT: tahun
S2	: Data Science dan Kecerdasan Buatan, USU	TAMAT: tahun

UCAPAN TERIMA KASIH

Puji syukur penulis panjatkan kepada Tuhan Yang Maha Esa atas rahmat dan karunia-Nya sehingga tesis yang berjudul “*Judul Tesis untuk Program Magister (S2) Data Science dan Kecerdasan Buatan*” ini dapat diselesaikan sebagai salah satu syarat memperoleh gelar Magister pada Program Studi S2 Data Science dan Kecerdasan Buatan, Fakultas Ilmu Komputer dan Teknologi Informasi, Universitas Sumatera Utara.

Penulis mengucapkan terima kasih kepada Komisi Pembimbing atas bimbingan, arahan, dan waktu yang telah diberikan selama proses penelitian dan penulisan tesis ini. Ucapan terima kasih juga disampaikan kepada seluruh dosen dan staf Program Studi S2 Data Science dan Kecerdasan Buatan, keluarga, serta rekan-rekan yang telah memberikan dukungan moril maupun materiil.

Penulis menyadari bahwa tesis ini masih memiliki kekurangan. Oleh karena itu, kritik dan saran yang membangun sangat diharapkan demi kesempurnaan penelitian selanjutnya.

Medan, 2026

Penulis

ABSTRAK

Perkembangan volume data yang pesat pada berbagai sektor menuntut pendekatan analisis yang mampu mengekstraksi pola dan pengetahuan secara otomatis dan akurat. Penelitian ini dilatarbelakangi oleh permasalahan tersebut, khususnya pada domain *[nama domain aplikasi, misalnya: deteksi anomali/klasifikasi teks/prediksi deret waktu]*, di mana pendekatan konvensional belum mampu memberikan hasil yang optimal dari segi akurasi maupun efisiensi komputasi. Penelitian ini bertujuan untuk merancang dan mengevaluasi sebuah pendekatan berbasis *[nama metode/algoritma, misalnya: pembelajaran mendalam, model probabilistik Bayesian, atau algoritma metaheuristik]* yang dapat meningkatkan kinerja pada permasalahan tersebut dibandingkan dengan metode *baseline*. Metodologi penelitian terdiri atas tahap pengumpulan dan praproses data, perancangan arsitektur model, pelatihan dan penyetelan hiperparameter, serta pengujian menggunakan skema validasi silang pada himpunan data *[nama/asal dataset]*. Kinerja model dievaluasi menggunakan metrik *[nama metrik, misalnya: akurasi, F1-score, RMSE]* dan dibandingkan dengan beberapa metode pembandingan dari literatur terkini. Hasil penelitian menunjukkan bahwa metode yang diusulkan memberikan peningkatan kinerja sebesar *[nilai/persentase]* dibandingkan dengan metode pembandingan, serta menunjukkan konsistensi hasil pada beberapa skenario pengujian. Temuan ini mengindikasikan bahwa pendekatan yang diusulkan berpotensi diterapkan pada permasalahan sejenis dalam skala yang lebih luas, sekaligus membuka peluang penelitian lanjutan terkait optimalisasi arsitektur model dan efisiensi komputasi.

Kata Kunci: kata kunci satu, kata kunci dua, kata kunci tiga, kata kunci empat.

ABSTRACT

The rapid growth of data volume across various sectors demands analytical approaches capable of automatically and accurately extracting patterns and knowledge. This research is motivated by that challenge, specifically within the domain of *[application domain, e.g., anomaly detection/text classification/time-series forecasting]*, where conventional approaches have not been able to deliver optimal results in terms of either accuracy or computational efficiency. This study aims to design and evaluate an approach based on *[method/algorithm name, e.g., deep learning, a Bayesian probabilistic model, or a metaheuristic algorithm]* that improves performance on the problem compared to a baseline method. The research methodology consists of data collection and preprocessing, model architecture design, training and hyperparameter tuning, and evaluation using a cross-validation scheme on the *[dataset name/source]* dataset. Model performance is evaluated using *[metric name, e.g., accuracy, F1-score, RMSE]* and compared against several state-of-the-art baseline methods. The results show that the proposed method achieves a performance improvement of *[value/percentage]* over the baseline methods, with consistent results across several testing scenarios. These findings indicate that the proposed approach has the potential to be applied to similar problems at a larger scale, while also opening opportunities for further research on model architecture optimization and computational efficiency.

Keywords: keyword one, keyword two, keyword three, keyword four.

DAFTAR ISI

1	PENDAHULUAN	2
1.1	Latar Belakang	2
1.1.1	Ruang lingkup permasalahan yang diteliti	2
1.2	Rumusan Masalah	2
1.3	Tujuan Penelitian	2
1.4	Manfaat Penelitian	3
1.5	Batasan Masalah	3
1.6	Sistematika Penulisan	3
2	TINJAUAN PUSTAKA	4
2.1	Penelitian Terdahulu	4
2.1.1	Ringkasan perbandingan penelitian terdahulu	4
2.2	Landasan Teori	4
2.2.1	Data science dan kecerdasan buatan	4
2.2.2	Metode yang relevan dengan penelitian ini	5
3	METODOLOGI PENELITIAN	6
3.1	Tahapan Penelitian	6
3.1.1	Kerangka kerja penelitian	6
3.2	Pengumpulan dan Praproses Data	6
3.3	Perancangan Model	6
3.3.1	Arsitektur model yang diusulkan	7
3.4	Skenario Pengujian	7
4	HASIL DAN PEMBAHASAN	8
4.1	Karakteristik Data	8
4.2	Hasil Pelatihan Model	8
4.3	Perbandingan Kinerja dengan Metode Pembanding	8
4.4	Pembahasan	9
5	KESIMPULAN DAN SARAN	10
5.1	Kesimpulan	10
5.2	Saran	10
A	Kode Program	12

B Data Tambahan

DAFTAR TABEL

2.1	Perbandingan penelitian terdahulu	4
4.1	Statistik ringkas himpunan data	8
4.2	Perbandingan kinerja model	9
B.1	Hasil pengujian lengkap per fold validasi silang	13

DAFTAR GAMBAR

4.1	Kurva <i>loss</i> pelatihan dan validasi (ilustrasi)	8
-----	--	---

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Bab Pendahuluan adalah bab terpendek dalam sebuah tesis dan merupakan bab yang paling penting. Satu-satunya tujuan penulisan bab ini adalah untuk memperkenalkan kepada pembaca tentang apa yang akan dipaparkan dalam tesis (Russell and Norvig, 2020).

Jelaskan lebih dahulu apa yang memotivasi penelitian yang akan dilakukan dan mengapa penelitian pada domain [*nama domain aplikasi*] tersebut perlu dilakukan. Paparkan secara sistematis gejala dan peristiwa yang menimbulkan permasalahan untuk diteliti, serta sebutkan penelitian-penelitian terdahulu yang relevan pada bidang yang akan diteliti (Goh et al., 2009).

1.1.1 Ruang lingkup permasalahan yang diteliti

Pada bagian ini penulis menjelaskan secara ringkas gambaran umum permasalahan penelitian, batasan awal ruang lingkupnya, serta perbedaan kontribusi penelitian ini dibandingkan dengan penelitian-penelitian sebelumnya yang telah disebutkan.

1.2 Rumusan Masalah

Rumusan masalah dapat ditulis dalam bentuk *research question* atau *problem statement*. Contoh rumusan masalah berbentuk pertanyaan penelitian:

1. Bagaimana pengaruh [*variabel/metode*] terhadap kinerja [*tugas/permasalahan*] pada [*domain aplikasi*]?
2. Bagaimana perbandingan kinerja pendekatan yang diusulkan terhadap metode *baseline* yang ada?

1.3 Tujuan Penelitian

Tujuan penelitian ini adalah untuk merancang, mengimplementasikan, dan mengevaluasi [*nama metode/algoritma*] guna menyelesaikan permasalahan yang telah dirumuskan pada Subbab 1.2.

1.4 Manfaat Penelitian

Manfaat yang diharapkan dari penelitian ini antara lain memberikan kontribusi teoritis pada bidang *[bidang keilmuan]* serta manfaat praktis berupa *[manfaat terapan, misalnya: purwarupa sistem, rekomendasi kebijakan, dsb.]*.

1.5 Batasan Masalah

Agar penelitian tetap fokus, ruang lingkup penelitian dibatasi pada hal-hal berikut:

1. Penelitian dibatasi pada disiplin ilmu data science dan kecerdasan buatan.
2. Data yang digunakan adalah *[nama/sumber dataset]*.
3. Evaluasi dilakukan menggunakan metrik *[nama metrik]* dan tidak mencakup *[hal yang dikecualikan]*.

1.6 Sistematika Penulisan

Sistematika penulisan tesis ini terdiri atas lima bab. Bab 1 berisi pendahuluan. Bab 2 membahas tinjauan pustaka dan landasan teori. Bab 3 menguraikan metodologi penelitian. Bab 4 menyajikan hasil dan pembahasan. Bab 5 berisi kesimpulan dan saran.

BAB 2

TINJAUAN PUSTAKA

2.1 Penelitian Terdahulu

Beberapa penelitian sebelumnya telah mengeksplorasi permasalahan serupa. Riedmiller (1994) mengusulkan algoritma *gradient descent* untuk meminimumkan *mean square error* (MSE) pada pelatihan jaringan saraf tiruan. Pendekatan ini kemudian dikembangkan lebih lanjut oleh beberapa peneliti lain (Calvert and Guan, 2005; Aaron et al., 2014).

Sementara itu, meskipun dianggap sebagai pendekatan yang relatif baru, algoritma *artificial bee colony* (ABC) telah digunakan secara luas dalam pencarian penyelesaian optimum (Shukran et al., 2011).

2.1.1 Ringkasan perbandingan penelitian terdahulu

Tabel 2.1 merangkum perbandingan pendekatan, dataset, dan metrik evaluasi yang digunakan pada beberapa penelitian terdahulu yang relevan dengan penelitian ini.

Tabel 2.1. Perbandingan penelitian terdahulu

Peneliti	Metode	Dataset	Metrik
Riedmiller (1994)	Gradient Descent	Sintetis	MSE
Shukran et al. (2011)	Artificial Bee Colony	Benchmark	Akurasi
Russell and Norvig (2020)	Pembelajaran Mesin	Umum	F1-score

2.2 Landasan Teori

2.2.1 Data science dan kecerdasan buatan

Data science merupakan bidang interdisipliner yang menggabungkan statistik, ilmu komputer, dan pengetahuan domain untuk mengekstraksi wawasan dari data. Kecerdasan buatan (*artificial intelligence*) adalah cabang ilmu komputer yang berfokus pada pembangunan sistem yang mampu meniru kemampuan kognitif manusia (Russell and Norvig, 2020).

2.2.2 Metode yang relevan dengan penelitian ini

Uraikan di sini teori dan konsep matematis dari metode/algorithm yang menjadi dasar pendekatan yang diusulkan pada Bab 3, misalnya arsitektur model pembelajaran mesin, fungsi *loss*, atau kerangka evaluasi yang digunakan.

BAB 3

METODOLOGI PENELITIAN

3.1 Tahapan Penelitian

Penelitian ini dilaksanakan melalui beberapa tahapan utama: (1) pengumpulan dan praproses data, (2) perancangan model, (3) pelatihan dan penyetelan hiperparameter, serta (4) pengujian dan evaluasi.

3.1.1 Kerangka kerja penelitian

Kerangka kerja penelitian secara umum mengikuti siklus iteratif antara perancangan model dan evaluasi hasil, hingga diperoleh konfigurasi model dengan kinerja terbaik pada data validasi.

3.2 Pengumpulan dan Praproses Data

Data yang digunakan dalam penelitian ini bersumber dari *[nama/sumber dataset]*. Tahap praproses meliputi pembersihan data, penanganan nilai hilang, normalisasi fitur, dan pembagian data menjadi himpunan latih, validasi, dan uji.

3.3 Perancangan Model

Diberikan sebuah himpunan data $D = \{(x_i, y_i)\}_{i=1}^n$, model yang diusulkan bertujuan mempelajari fungsi $f_\theta : \mathcal{X} \rightarrow \mathcal{Y}$ yang meminimumkan fungsi kerugian (*loss function*) berikut:

$$\mathcal{L}(\theta) = \frac{1}{n} \sum_{i=1}^n \ell(f_\theta(x_i), y_i) + \lambda \|\theta\|_2^2 \quad (3.1)$$

dengan $\ell(\cdot, \cdot)$ adalah fungsi kerugian dasar, θ adalah parameter model, dan λ adalah koefisien regularisasi.

Parameter optimal θ^* diperoleh melalui optimisasi berbasis gradien sebagaimana pada Persamaan (3.2):

$$\theta^{(t+1)} = \theta^{(t)} - \eta \nabla_\theta \mathcal{L}(\theta^{(t)}) \quad (3.2)$$

dengan η adalah laju pembelajaran (*learning rate*) dan t adalah indeks iterasi.

3.3.1 *Arsitektur model yang diusulkan*

Jelaskan di sini arsitektur model secara rinci, termasuk lapisan/komponen penyusun, fungsi aktivasi yang digunakan, serta alasan pemilihan arsitektur tersebut dikaitkan dengan tinjauan pustaka pada Bab 2.

3.4 Skenario Pengujian

Evaluasi kinerja model dilakukan menggunakan skema validasi silang *k-fold* dengan metrik evaluasi berupa [*nama metrik*]. Hasil pengujian dibandingkan dengan beberapa metode *baseline* sebagaimana diuraikan pada Bab 2.

BAB 4

HASIL DAN PEMBAHASAN

4.1 Karakteristik Data

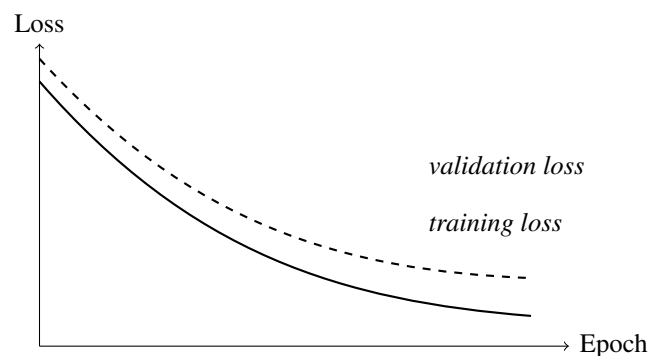
Bagian ini memaparkan karakteristik data hasil praproses yang digunakan pada tahap pelatihan dan pengujian model, sebagaimana diuraikan pada Subbab 3.2.

Tabel 4.1. Statistik ringkas himpunan data

Himpunan	Jumlah Sampel	Jumlah Fitur	Proporsi (%)
Latih	0	0	70
Validasi	0	0	15
Uji	0	0	15

4.2 Hasil Pelatihan Model

Gambar 4.1 memperlihatkan kurva *loss* pelatihan dan validasi terhadap jumlah iterasi (*epoch*). Kurva ini digunakan untuk memantau kekonvergenan model serta mendeteksi indikasi *overfitting*.



Gambar 4.1. Kurva *loss* pelatihan dan validasi (ilustrasi)

4.3 Perbandingan Kinerja dengan Metode Pembanding

Tabel 4.2 menyajikan perbandingan kinerja pendekatan yang diusulkan dengan beberapa metode *baseline* yang telah disebutkan pada Bab 2 (Russell and Norvig, 2020; Shukran et al., 2011).

Tabel 4.2. Perbandingan kinerja model

Metode	Akurasi	F1-score	Waktu Komputasi (s)
<i>Baseline 1</i>	0.00	0.00	0.00
<i>Baseline 2</i>	0.00	0.00	0.00
Metode yang Diusulkan	0.00	0.00	0.00

4.4 Pembahasan

Hasil pada Tabel 4.2 menunjukkan bahwa metode yang diusulkan memberikan kinerja *[lebih baik/setara/dst.]* dibandingkan metode pembanding. Diskusikan di sini keterkaitan hasil dengan rumusan masalah pada Subbab 1.2 dan tinjauan pustaka pada Bab 2, termasuk kemungkinan penjelasan atas hasil yang diperoleh serta keterbatasannya.

BAB 5

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil penelitian dan pembahasan pada Bab 4, dapat disimpulkan bahwa:

1. Metode yang diusulkan pada penelitian ini mampu menjawab tujuan penelitian sebagaimana dirumuskan pada Subbab 1.3.
2. Pendekatan yang diusulkan menunjukkan kinerja [*lebih baik/setara*] dibandingkan metode pembandingan pada metrik evaluasi yang digunakan.

5.2 Saran

Beberapa saran untuk penelitian selanjutnya antara lain:

1. Mengeksplorasi [*arsitektur/metode lain*] untuk meningkatkan kinerja lebih lanjut.
2. Menguji generalisasi pendekatan yang diusulkan pada *dataset* lain di luar cakupan penelitian ini.
3. Mengkaji efisiensi komputasi model untuk kebutuhan penerapan pada skala produksi.

DAFTAR PUSTAKA

- Aaron, C. et al. (2014). Adaptive distributed neural learning for large-scale systems. In *Proceedings of the International Conference on Machine Learning*, pages 112–120.
- Calvert, K. and Guan, J. (2005). A distributed autonomous neuro-gen learning engine. *Journal of Computational Intelligence*, 2:45–58.
- Goh, Y. M., Giess, M., McMahon, C., and Liu, C. (2009). From faceted classification to knowledge discovery of semi-structured text records. In Abraham, A. et al., editors, *Foundation of Computational Intelligence*, volume 6, pages 151–169.
- Riedmiller, M. (1994). Advanced supervised learning in multi-layer perceptrons – from backpropagation to adaptive learning algorithms. *Computer Standards & Interfaces*, 16(3):265–278.
- Russell, S. and Norvig, P. (2020). *Artificial Intelligence: A Modern Approach*. Pearson, New Jersey, 4th edition.
- Shukran, M. A. M. et al. (2011). Artificial bee colony algorithm for optimum solution search. *International Journal of Computer Applications*, 30(5):1–6.

LAMPIRAN A

KODE PROGRAM

Lampiran ini memuat *source code* yang dibangun sebagai hasil implementasi model yang dikemukakan pada Bab 3.

Listing A.1: Contoh cuplikan kode pelatihan model

```
def train(model, dataloader, optimizer, loss_fn, epochs):  
    for epoch in range(epochs):  
        for x_batch, y_batch in dataloader:  
            optimizer.zero_grad()  
            y_pred = model(x_batch)  
            loss = loss_fn(y_pred, y_batch)  
            loss.backward()  
            optimizer.step()  
    return model
```

LAMPIRAN B

DATA TAMBAHAN

Lampiran ini memuat data mentah, tabel hasil pengujian lengkap, atau materi pendukung lain yang tidak dimuat pada teks utama, sebagaimana diuraikan pada Subbab 6.4 Pedoman Penulisan Tesis.

Tabel B.1. Hasil pengujian lengkap per fold validasi silang

Fold	Akurasi	F1-score	Waktu (s)
1	0.00	0.00	0.00
2	0.00	0.00	0.00
3	0.00	0.00	0.00
4	0.00	0.00	0.00
5	0.00	0.00	0.00