



SKRIPSI

**KOMPARASI KINERJA *DIETCLASSIFIER*
DALAM KLASIFIKASI *INTENT*
DATA SEJARAH DEWI DURGA
DENGAN MENGGUNAKAN *FRAMEWORK RASA***

DEA PUSPITA ANGGRAENI

NPM 21082010029

DOSEN PEMBIMBING

Eka Dyar Wahyuni, S.Kom., M.Kom.

Tri Lathif Mardi Suryanto, S.Kom., M.T.

**KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI SISTEM INFORMASI
SURABAYA**

2025



SKRIPSI

**KOMPARASI KINERJA *DIETCLASSIFIER*
DALAM KLASIFIKASI *INTENT*
DATA SEJARAH DEWI DURGA
DENGAN MENGGUNAKAN *FRAMEWORK RASA***

DEA PUSPITA ANGGRAENI

NPM 21082010029

DOSEN PEMBIMBING

Eka Dyar Wahyuni, S.Kom., M.Kom.

Tri Lathif Mardi Suryanto, S.Kom., M.T.

**KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI SISTEM INFORMASI
SURABAYA**

2025

Halaman ini sengaja dikosongkan

LEMBAR PENGESAHAN

KOMPARASI KINERJA *DIETCLASSIFIER* DALAM KLASIFIKASI *INTENT* DATA SEJARAH DEWI DURGA DENGAN MENGGUNAKAN *FRAMEWORK RASA*

Oleh :

DEA PUSPITA ANGGRAENI
NPM. 21082010029

Telah dipertahankan dihadapan dan diterima oleh Tim Penguji Skripsi Prodi Sistem Informasi Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jawa Timur Pada tanggal 10 Juli 2025

Menyetujui

Eka Dyar Wahyuni, S.Kom., M.Kom.
NIP. 198412012021212005

..... (Pembimbing I)

Tri Lathif Mardi Suryanto, S.Kom., M.T.
NIP. 198902252021211001

..... (Pembimbing II)

Agung Brastama Putra, S.Kom., M.Kom.
NIP. 198511242021211003

..... (Ketua Penguji)

Abdul Rezha Efrat Najaf, S.Kom., M.Kom.
NIP. 199409292022031008

..... (Anggota Penguji II)

Prasasti Karunia F.A., S.Kom., M.Kom., M.IM
NIP. 199707042024062001

..... (Anggota Penguji III)

Mengetahui,
Dekan Fakultas Ilmu Komputer


Prof. Dr. Ir. Novirina Hendrasarie, MT
NIP. 19681126 199403 2 001

Halaman ini sengaja dikosongkan

LEMBAR PERSETUJUAN

KOMPARASI KINERJA *DIETCLASSIFIER* DALAM KLASIFIKASI *INTENT DATA* SEJARAH DEWI DURGA DENGAN MENGUNAKAN *FRAMEWORK RASA*

Oleh :

DEA PUSPITA ANGGRAENI

NPM. 21082010029

Mengetahui,

Koordinator Program Studi Sistem Informasi

Fakultas Ilmu Komputer



Agung Brastama Putra, S.Kom., M.Kom.
NIP. 198511242021211003

Halaman ini sengaja dikosongkan

SURAT PERNYATAAN BEBAS PLAGIASI

Yang bertandatangan di bawah ini:

Nama Mahasiswa : Dea Puspita Anggraeni
NPM : 21082010029
Program : Sarjana (S1)
Program Studi : Sistem Informasi
Fakultas : Ilmu Komputer

Menyatakan bahwa dalam dokumen ilmiah Skripsi ini tidak terdapat bagian dari karya ilmiah lain yang telah diajukan memperoleh gelar akademik di suatu lembaga Pendidikan Tinggi, dan juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang/lembaga lain, kecuali yang secara tertulis disitasi dalam dokumen ini dan disebutkan secara lengkap dalam daftar pustaka.

Dan saya menyatakan bahwa dokumen ilmiah ini bebas dari unsur-unsur plagiasi. Apabila dikemudian hari ditemukan indikasi plagiat pada Skripsi ini, saya bersedia menerima sanksi sesuai dengan peraturan perundang-undangan yang berlaku.

Demikian surat pernyataan ini saya buat dengan sesungguhnya tanpa ada paksaan dari siapapun juga dan untuk dipergunakan sebagaimana mestinya.



Surabaya, 22 Juli 2025
Yang Membuat Pernyataan,



Dea Puspita Anggraeni
NPM. 21082010029

Halaman ini sengaja dikosongkan

ABSTRAK

Nama Mahasiswa / NPM : Dea Puspita Anggraeni / 21082010029

Judul Skripsi : Komparasi Kinerja *DIETClassifier*
dalam Klasifikasi *Intent* Data Sejarah Dewi Durga
dengan Menggunakan *Framework Rasa*

Dosen Pembimbing : 1. Eka Dyar Wahyuni, S.Kom., M.Kom.
2. Tri Lathif Mardi Suryanto, S.Kom., M.T.

Data sejarah memiliki struktur yang tidak teratur dan mengandung makna simbolik yang mendalam, sehingga memerlukan pendekatan yang tepat untuk memproses dan mengklasifikasikannya. Penelitian ini menerapkan model *DIETClassifier* pada data sejarah Dewi Durga menggunakan framework Rasa untuk mengklasifikasikan intent dalam teks naratif. Pendekatan yang digunakan adalah CRISP-DM, dimulai dengan studi literatur, diikuti dengan pemahaman dan persiapan data. Data yang telah diproses dengan NLP kemudian diberi label *intent* dan dibagi menjadi data latih dan data uji. Model *DIETClassifier* dilatih dengan sembilan konfigurasi *hyperparameter* berbeda. Konfigurasi terbaik, yang menggunakan 100 *epoch*, *constrain similarity* bernilai *true*, *fine-tune* bernilai *true*, *batch size* 16, *learning rate* 0.001, *optimizer adamax*, dan *dropout rate* 0.2, menghasilkan akurasi 99,63% dan F1-score 99,57% pada data uji, yang menunjukkan performa terbaik model dalam klasifikasi *intent*. Evaluasi dilakukan menggunakan *confusion matrix* akurasi, presisi, *recall*, dan *F1-score* untuk mengukur kinerja model. Setelah pelatihan, model dilakukan *deployment* secara lokal dan diintegrasikan ke dalam aplikasi berbasis web. Penelitian ini memberikan kontribusi pada pengembangan *chatbot* berbasis NLP untuk pelestarian dan penyajian informasi sejarah serta budaya secara lebih efisien, sekaligus membuka peluang pengembangan lebih lanjut dalam analisis teks budaya menggunakan teknologi NLP.

Kata Kunci : *DIETClassifier*, Klasifikasi *Intent*, Dewi Durga, *Framework Rasa*, CRISP-DM, *Confusion Matrix*

Halaman ini sengaja dikosongkan

ABSTRACT

Student Name / NPM : Dea Puspita Anggraeni / 21082010029

Thesis Title : Comparing the Performance of *DIETClassifier*
in *Intent* Classification of Goddess Durga
History Data Using the Rasa *Framework*

Advisors : 1. Eka Dyar Wahyuni, S.Kom., M.Kom.
2. Tri Lathif Mardi Suryanto, S.Kom., M.T.

Historical data has an irregular structure and contains deep symbolic meaning, requiring an appropriate approach to process and classify it. This research applies the *DIETClassifier* model to historical data of Dewi Durga using the Rasa framework to classify intents in narrative texts. The approach used is CRISP-DM, starting with literature study, followed by data understanding and preparation. The data processed with NLP is then labeled with *intents* and divided into training and testing data. The *DIETClassifier* model is trained with nine different *hyperparameter* configurations. The best configuration, which uses 100 *epochs*, *constrain similarity* set to *true*, *fine-tune* set to *true*, *batch size* 16, *learning rate* 0.001, *adamax optimizer*, and *dropout rate* 0.2, achieves 99.63% accuracy and 99.57% F1-score on test data, demonstrating the model's best performance in intent classification. Evaluation is conducted using *confusion matrix*, accuracy, precision, *recall*, and *F1-score* to measure model performance. After training, the model is deployed locally and integrated into a web-based application. This research contributes to the development of NLP-based *chatbots* for preserving and presenting historical and cultural information more efficiently, while opening opportunities for further development in cultural text analysis using NLP technology.

Keywords : DIETClassifier, Intent Classification, Dewi Durga, Rasa Framework, CRISP-DM, Confusion Matrix

Halaman ini sengaja dikosongkan

KATA PENGANTAR

Rasa syukur yang tidak terhingga penulis panjatkan kepada Allah SWT yang telah memberikan rahmat-Nya dan karunia-Nya sehingga penulis diberikan kekuatan untuk menyelesaikan semua proses sulit dalam penyusunan skripsi ini yang berjudul **“Komparasi Kinerja *DIETClassifier* dalam Klasifikasi *Intent Data Sejarah Dewi Durga* dengan Menggunakan *Framework Rasa*”** dengan baik. Adapun maksud penulisan skripsi ini ditujukan untuk memenuhi syarat memperoleh gelar Sarjana Komputer di kampus tercinta Universitas Pembangunan Nasional “Veteran” Jawa Timur.

Pembuatan Skripsi ini penuh tantangan yang tidak terlepas dari jerih payah dan bantuan banyak pihak. Oleh karena itu, penulis mengucapkan banyak terima kasih kepada :

1. Ayah dan Ibu, yang selalu ada untuk penulis, mendengarkan setiap keluh kesah, memanjakan, mendoakan, serta memberi dukungan materiil. Dengan dukungan mereka penulis bisa melewati masa-masa penuh tekanan ini.
2. Ibu Eka Dyar Wahyuni, S.Kom., M.Kom. dan Bapak Tri Lathif Mardi Suryanto, S.Kom., M.T., selaku dosen pembimbing, yang dengan sabar memberikan waktu, arahan, dan bimbingan yang sangat berharga selama proses penyusunan skripsi.
3. Kakak, yang selalu mengapresiasi, memberikan semangat dan asupan makanan yang enak, serta menjadi tempat berbagi cerita mengenai kepenatan selama proses mengerjakan skripsi.
4. Para sahabat Tiara, Desi, dan Dona, yang selalu memotivasi, memberikan dukungan, dan dengan sabar mendengarkan curhatan penulis.

5. Teman-teman dan partner kuliah, Carissa Renatasari dan Devilia Dwi Candra, yang menemani masa-masa kuliah dari awal hingga akhir. Dukungan dan kehadiran yang diberikan sangat berarti dalam menyelesaikan studi ini.

Penulis menyadari bahwa skripsi ini masih memiliki kekurangan dalam penyusunannya, oleh karena itu, penulis dengan terbuka menerima setiap kritik dan saran yang membangun. Semoga skripsi ini dapat menjadi pembelajaran dan referensi yang bermanfaat, baik bagi penulis maupun pembaca.

Surabaya, 01 Juli 2025

Penulis

DAFTAR ISI

LEMBAR PENGESAHAN	iii
LEMBAR PERSETUJUAN	v
SURAT PERNYATAAN BEBAS PLAGIASI.....	vii
ABSTRAK	ix
ABSTRACT	xi
KATA PENGANTAR.....	xiii
DAFTAR ISI.....	xv
DAFTAR GAMBAR.....	xix
DAFTAR TABEL	xxi
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Batasan Masalah.....	5
1.4 Tujuan.....	5
1.5 Sistematika Penulisan.....	6
BAB II TINJAUAN PUSTAKA.....	9
2.1 Penelitian Terdahulu.....	9
2.2 Dasar Teori	13
2.2.1 Data Durga.....	13
2.2.2 <i>Python</i>	19
2.2.3 <i>NumPy</i>	21
2.2.4 Sastrawi.....	23
2.2.5 <i>Chatbot</i>	24
2.2.6 <i>NLP (Natural Language Processing)</i>	26

2.2.7 <i>Framework Rasa</i>	28
2.2.8 <i>WhitespaceTokenizer</i>	30
2.2.9 <i>CountVectorsFeaturizer</i>	31
2.2.10 <i>DIETClassifier</i>	31
2.2.11 <i>FallbackClassifier</i>	32
2.2.12 <i>Confusion Matrix</i>	33
2.2.13 <i>Flask</i>	35
BAB III METODOLOGI PENELITIAN	37
3.1 Studi Literatur	38
3.2 Pemahaman Data	39
3.2.1 Pengumpulan Data	39
3.2.2 Pembersihan Data	40
3.3 Persiapan Data	42
3.3.1 Filter Data Durga atau Dewi Durga	42
3.3.2 Pemberian Label <i>Intent</i>	43
3.3.3 Membagi Data <i>Train</i> dan Data <i>Test</i>	43
3.3.4 Preproses <i>NLP</i>	43
3.3.5 Penyusunan File Rasa	46
3.4 Pembuatan Model dan Skenario	49
3.5 Evaluasi	54
3.6 <i>Deployment</i>	55
BAB IV HASIL DAN PEMBAHASAN	57
4.1 Pemahaman Data	57
4.1.1 Pengumpulan Data	57
4.1.2 Pembersihan Data	58
4.2 Persiapan Data	63

4.2.1 Filter Data Durga atau Dewi Durga	63
4.2.2 Pemberian Label <i>Intent</i>	64
4.2.3 Membagi Data <i>Train</i> dan Data <i>Test</i>	65
4.2.4 Preproses <i>NLP</i>	66
4.2.5 Penyusunan File Rasa	71
4.3 Pembuatan Model dan Skenario	75
4.4 Evaluasi	79
4.5 <i>Deployment</i>	85
BAB V_PENUTUP	89
5.1 Kesimpulan	89
5.2 Saran	90
DAFTAR PUSTAKA	93

Halaman ini sengaja dikosongkan

DAFTAR GAMBAR

Gambar 2.1 Arsitektur Percakapan <i>Framework</i> Rasa.....	29
Gambar 3.1 Tahapan Penelitian	38
Gambar 3.2 Bentuk Data Asli Sejarah Durga	40
Gambar 3.3 Preproses <i>NLP</i>	44
Gambar 3.4 Ilustrasi <i>Case Folding</i>	44
Gambar 3.5 Ilustrasi <i>Tokenizing</i>	45
Gambar 3.6 Ilustrasi <i>Remove Stopwords</i>	45
Gambar 3.7 Ilustrasi <i>Stemming</i>	46
Gambar 3.8 Pemodelan Data Pada <i>Framework</i> Rasa	46
Gambar 3.9 Ilustrasi <i>nlu.yml</i>	47
Gambar 3.10 Ilustrasi <i>domain.yml</i>	48
Gambar 3.11 Ilustrasi <i>stories.yml</i>	48
Gambar 3.12 Ilustrasi <i>config.yml</i>	49
Gambar 3.13 Tahap Deployment <i>Framework</i> Rasa.....	55
Gambar 3.14 Desain Awal <i>User Interface Chatbot</i>	56
Gambar 4.1 Data Asli Sejarah Durga.....	57
Gambar 4.2 Hasil Pengecekan Duplikasi Awal Data Sejarah Durga.....	59
Gambar 4.3 <i>Source Code</i> Pengecekan Duplikasi Tiap Kolom	61
Gambar 4.4 Hasil Pengecekan Duplikasi Tiap Kolom	61
Gambar 4.5 <i>Source Code</i> Penggeseran Nilai Kosong.....	62
Gambar 4.6 Hasil Penggeseran Nilai Kosong.....	63
Gambar 4.7 <i>Source Code</i> Filter Data Durga	63
Gambar 4.8 Hasil Pemberian Label <i>Intent</i>	65
Gambar 4.9 <i>Source Code</i> Pembagian Data <i>Train</i> dan Data <i>Test</i>	66
Gambar 4.10 Pemrosesan <i>Case Folding</i> Data <i>Train</i>	67
Gambar 4.11 Pemrosesan <i>Case Folding</i> Data <i>Test</i>	67
Gambar 4.12 Pemrosesan <i>Tokenizing</i> Data <i>Train</i>	68
Gambar 4.13 Pemrosesan <i>Tokenizing</i> Data <i>Test</i>	68
Gambar 4.14 Pemrosesan <i>Remove Stopwords</i> Data <i>Train</i>	69

Gambar 4.15 Pemrosesan <i>Remove Stopwords</i> Data <i>Test</i>	69
Gambar 4.16 Pemrosesan <i>Stemming</i> Data <i>Train</i>	69
Gambar 4.17 Pemrosesan <i>Stemming</i> Data <i>Test</i>	70
Gambar 4.18 Hasil Preproses <i>NLP</i> Data <i>Train</i>	70
Gambar 4.19 Hasil Preproses <i>NLP</i> Data <i>Test</i>	71
Gambar 4.20 Proses Pembuatan File <i>nlu.yml</i>	72
Gambar 4.21 Struktur File <i>nlu.yml</i>	72
Gambar 4.22 Proses Pembuatan File <i>domain.yml</i>	73
Gambar 4.23 Struktur File <i>domain.yml</i>	73
Gambar 4.24 Proses Pembuatan File <i>stories.yml</i>	74
Gambar 4.25 Struktur File <i>stories.yml</i>	75
Gambar 4.26 Susunan <i>Pipeline</i> <i>config.yml</i>	76
Gambar 4.27 Susunan <i>Policies</i> <i>config.yml</i>	77
Gambar 4.28 Proses Memodelkan Data Menggunakan Rasa <i>Train</i>	79
Gambar 4.29 Proses Evaluasi Menggunakan Rasa <i>Test</i>	79
Gambar 4.30 Proses Perbandingan Metrik Evaluasi dari Data <i>Test</i>	80
Gambar 4.31 Hasil Metrik Evaluasi Data <i>Test</i>	80
Gambar 4.32 Proses Perbandingan Metrik Evaluasi dari Data <i>Train</i>	81
Gambar 4.33 Hasil Metrik Hasil Evaluasi Data <i>Train</i>	82
Gambar 4.34 Gap Akurasi	82
Gambar 4.35 Kurva Akurasi (<i>epoch_i_acc</i>) Config 1	83
Gambar 4.36 Kurva <i>Loss</i> (<i>epoch_t_loss</i>) Config 1	84
Gambar 4.37 <i>Source Code Front End Chatbot</i>	85
Gambar 4.38 Konfigurasi <i>Endpoint</i> Rasa	86
Gambar 4.39 <i>User Interface Chatbot</i>	87

DAFTAR TABEL

Tabel 2. 1 Penelitian Terdahulu 1	9
Tabel 2. 2 Penelitian Terdahulu 2	10
Tabel 2. 3 Penelitian Terdahulu 3	11
Tabel 2. 4 Penelitian Terdahulu 4	12
Tabel 2. 5 Bentuk Data Durga.....	14
Tabel 2. 6 Confussion Matrix.....	33
Tabel 3. 1 Skenario Hyperparameter DIETClassifier	52

Halaman ini sengaja dikosongkan