

**KLASIFIKASI MULTI-LABEL TOPIK AL-QUR'AN PADA TEKS
TERJEMAHAN BAHASA INGGRIS MENGGUNAKAN BERT DAN BI-
LSTM DENGAN FITUR LINGUISTIK**

Skripsi



Disusun Oleh:

**AGUSTINA LATIFATUL MUTMAINNAH
(09050622079)**

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SUNAN AMPEL
2026**

PERNYATAAN KEASLIAN

Saya yang bertanda tangan dibawah ini

Nama : Agustina Latifatul Mutmainnah
Nim : 09050622079
Program Studi : Sistem Informasi
Angkatan : 2022

Menyatakan bahwa saya tidak melakukan plagiat dalam penulisan skripsi saya dengan judul “KLASIFIKASI MULTI-LABEL TOPIK AL-QUR'AN PADA TEKS TERJEMAHAN BAHASA INGGRIS MENGGUNAKAN BERT DAN BI-LSTM DENGAN FITUR LINGUISTIK”. Apabila suatu saat nanti terbukti saya melakukan tindakan plagiat, maka saya bersedia menerima sanksi yang telah ditetapkan.

Demikian pernyataan keaslian ini saya buat dengan sebenar-benarnya.

Surabaya, 11 Juni 2026

Yang menyatakan,



(Agustina Latifatul Mutmainnah)
NIM.09050622079

LEMBAR PERSETUJUAN PEMBIMBING

Skripsi oleh

NAMA : AGUSTINA LATIFATUL MUTMAINNAH

NIM : 09050622079

JUDUL : KLASIFIKASI MULTILABEL TOPIK AL-QUR'AN DALAM
TERJEMAHAN BAHASA INGGRIS MENGGUNAKAN
BERT, BI-LSTM DAN FITUR LINGUISTIK

Ini telah diperiksa dan disetujui untuk diujikan.

Surabaya, 05 Juni 2026

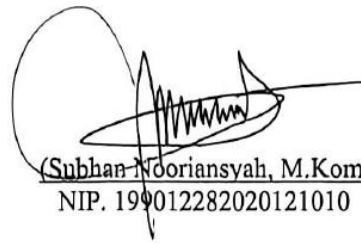
Dosen Pembimbing 1



(Dr. Eng. Anang Kunaefi, M. Kom)

NIP. 197911132014031001

Dosen Pembimbing 2



(Subhan Nooriansyah, M. Kom)

NIP. 199012282020121010

PENGESAHAN TIM PENGUJI SKRIPSI

Skripsi Agustina Latifatul Mutmainnah ini telah dipertahankan
di depan tim penguji skripsi
di Surabaya, 18 Juni 2026

Mengesahkan,
Dewan Penguji

Penguji I



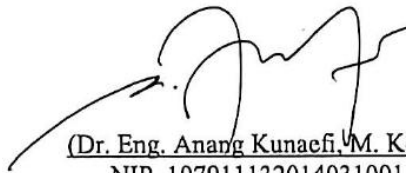
(Bayu Adhi Nugroho, Ph.D)
NIP. 197905182014031001

Penguji II



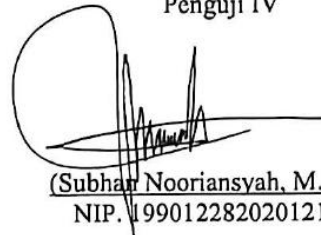
(Dwi Roflawati, M. T)
NIP. 197909272014032001

Penguji III



(Dr. Eng. Anang Kunaefi, M. Kom)
NIP. 197911132014031001

Penguji IV



(Subhan Nooriansyah, M. Kom)
NIP. 199012282020121010

Mengetahui,

Surabaya, Fakultas Sains dan Teknologi
Jl. Sunan Ampel Surabaya



(Saepul Hamdani, M.Pd)
NIP. 196507312000031002

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademika UIN Sunan Ampel Surabaya, yang bertanda tangan di bawah ini, saya:

Nama : Agustina Latifatul Mutmainnah
NIM : 09050622079
Fakultas/Jurusan : Sistem Informasi
E-mail address : agustinalatifa11@gmail.com

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Perpustakaan UIN Sunan Ampel Surabaya, Hak Bebas Royalti Non-Eksklusif atas karya ilmiah :

☒ Sekripsi ☐ Tesis ☐ Desertasi ☐ Lain-lain (.....)
yang berjudul :

KLASIFIKASI MULTILABEL TOPIK AL-QUR'AN DALAM TERJEMAHAN BAHASA INGGRIS

MENGGUNAKAN BERT DAN BI-LSTM DENGAN FITUR LINGUISTIK

beserta perangkat yang diperlukan (bila ada). Dengan Hak Bebas Royalti Non-Eksklusif ini Perpustakaan UIN Sunan Ampel Surabaya berhak menyimpan, mengalih-media/format-kan, mengelolanya dalam bentuk pangkalan data (database), mendistribusikannya, dan menampilkan/mempublikasikannya di Internet atau media lain secara *fulltext* untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan atau penerbit yang bersangkutan.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Perpustakaan UIN Sunan Ampel Surabaya, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini yang saya buat dengan sebenarnya.

Surabaya,

Penulis



(Agustina Latifatul Mutmainnah)

ABSTRAK

KLASIFIKASI MULTI-LABEL TOPIK AL-QUR'AN PADA TEKS TERJEMAHAN BAHASA INGGRIS MENGGUNAKAN BERT DAN BI- LSTM DENGAN FITUR LINGUISTIK

Oleh:

Agustina Latifatul Mutmainnah

Al-Qur'an terdiri atas 6.236 ayat yang terdiri dari beberapa tema, dengan kemungkinan satu ayat dapat memuat beberapa tema, sehingga memiliki karakteristik klasifikasi multilabel. Karakteristik tersebut memerlukan model klasifikasi yang mampu memahami makna kontekstual setiap ayat secara lebih baik. Namun, penelitian-penelitian sebelumnya mengenai klasifikasi multilabel topik Al-Qur'an umumnya masih memanfaatkan *static word embedding*, yang memiliki keterbatasan dalam merepresentasikan makna kata sesuai dengan konteks penggunaannya. Oleh karena itu, penelitian ini bertujuan untuk meningkatkan performa klasifikasi multilabel topik Al-Qur'an pada terjemahan bahasa Inggris melalui integrasi *contextual embedding Bidirectional Encoder Representations from Transformers* (BERT), model *Bidirectional Long Short-Term Memory* (Bi-LSTM), serta fitur linguistik berupa *Part-of-Speech* (POS) Tagging dan *Named Entity Recognition* (NER). Penelitian ini dilakukan melalui tiga skenario eksperimen yang disusun secara bertahap untuk mengevaluasi kontribusi setiap komponen terhadap kinerja model. Evaluasi model menggunakan metrik Macro F1-score dan Hamming Loss, sedangkan hasil prediksi divalidasi oleh tiga pakar menggunakan koefisien Fleiss' Kappa dan Jaccard Similarity. Hasil penelitian menunjukkan bahwa kombinasi BERT, Bi-LSTM, POS Tagging, dan NER menghasilkan performa terbaik dengan nilai Macro F1-score sebesar 0,4835 dan Hamming Loss sebesar 0,0919, yang menunjukkan bahwa penambahan fitur linguistik mampu meningkatkan kemampuan model dalam mengidentifikasi berbagai topik pada setiap ayat. Validasi pakar menghasilkan nilai Fleiss' Kappa sebesar 0,3849 dan Jaccard Similarity sebesar 88,27%, yang menunjukkan bahwa hasil klasifikasi memiliki tingkat kesesuaian yang baik dengan penilaian pakar. Meskipun demikian, performa model secara keseluruhan masih belum optimal karena dipengaruhi oleh ketidakseimbangan distribusi label (*label imbalance*) dan karakteristik klasifikasi multilabel yang melibatkan banyak label dalam satu ayat. Oleh karena itu, penelitian selanjutnya disarankan untuk menerapkan strategi penanganan ketidakseimbangan data, mengeksplorasi arsitektur pembelajaran yang lebih adaptif, serta mengoptimalkan kualitas dan distribusi dataset guna meningkatkan performa klasifikasi multilabel topik Al-Qur'an.

Kata Kunci: Klasifikasi Multilabel, Topik Al-Qur'an, BERT, Bi-LSTM, POS Tagging, NER

ABSTRACT

MULTI-LABEL CLASSIFICATION OF AL-QUR'AN TOPIC IN ENGLISH TRANSLATION USING BERT AND BI-LSTM WITH LINGUISTIC FEATURES

By:

Agustina Latifatul Mutmainnahex

The Qur'an consists of 6,236 verses covering a wide range of themes, with a single verse often addressing multiple topics simultaneously, thereby exhibiting the characteristics of a multilabel classification problem. Such characteristics require a classification model capable of capturing the contextual meaning of each verse effectively. However, previous studies on multilabel Qur'anic topic classification have predominantly relied on static word embeddings, which are limited in representing word meanings according to their contextual usage. Therefore, this study aims to improve the performance of multilabel topic classification on the English translation of the Qur'an by integrating Bidirectional Encoder Representations from Transformers (BERT) as a contextual embedding model, Bidirectional Long Short-Term Memory (Bi-LSTM), and linguistic features, namely Part-of-Speech (POS) Tagging and Named Entity Recognition (NER). The proposed approach was evaluated through three progressive experimental scenarios designed to assess the contribution of each component to the overall model performance. Model performance was evaluated using the Macro F1-score and Hamming Loss, while the prediction results were validated by three domain experts using Fleiss' Kappa and Jaccard Similarity. The experimental results demonstrate that the combination of BERT, Bi-LSTM, POS Tagging, and NER achieved the best performance, yielding a Macro F1-score of 0.4835 and a Hamming Loss of 0.0919. These findings indicate that the incorporation of linguistic features improves the model's ability to identify multiple topics within a single verse. Furthermore, expert validation resulted in a Fleiss' Kappa coefficient of 0.3849 and a Jaccard Similarity of 88.27%, indicating that the predicted labels exhibit good agreement with expert judgments. Nevertheless, the overall performance of the proposed model has not yet reached its optimal level due to the severe label imbalance and the inherent characteristics of multilabel classification, where each verse may be associated with multiple labels. Therefore, future research is recommended to investigate more effective strategies for handling imbalanced data, explore more adaptive deep learning architectures, and optimize dataset quality and label distribution to further improve the performance of multilabel Qur'anic topic classification.

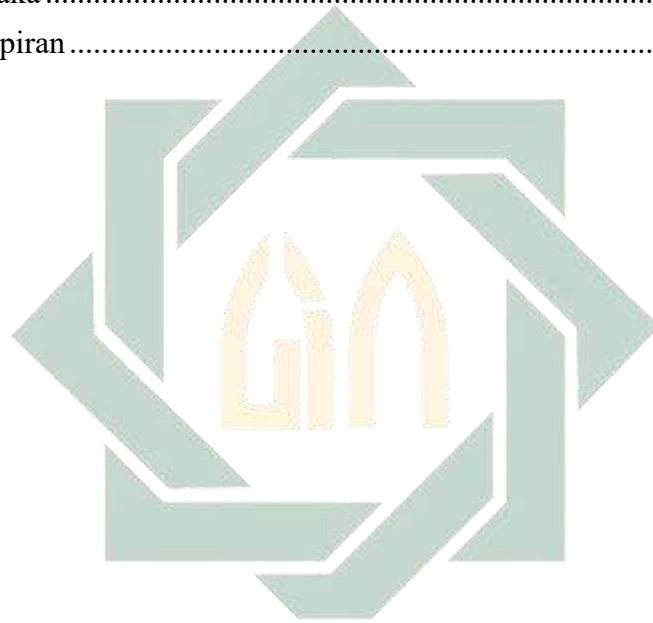
Key Words: *Multilable Classification, Al-Qur'an Topics, BERT, Bi-LSTM, POS Tagging, NER*

DAFTAR ISI

LEMBAR PERSETUJUAN PEMBIMBING	ii
PENGESAHAN TIM PENGUJI SKRIPSI.....	iii
LEMBAR PERNYATAAN PUBLIKASI	iv
PERNYATAAN KEASLIAN.....	v
HALAMAN PERSEMBAHAN	vi
KATA PENGANTAR	viii
ABSTRAK	x
ABSTRACT.....	xi
DAFTAR ISI.....	xii
DAFTAR GAMBAR	xv
DAFTAR TABEL.....	xvii
DAFTAR LAMPIRAN.....	xix
BAB I	20
1.1 Latar Belakang	20
1.2 Rumusan Masalah	23
1.3 Batasan Masalah.....	23
1.4 Tujuan Penelitian	23
1.5 Manfaat Penelitian	24
BAB II.....	25
2.1 Teori Penelitian Terdahulu.....	25
2.2 Teori Dasar.....	28
2.2.1 NLP.....	28
2.2.2 Deep learning.....	29
2.2.3 Teks Mining.....	33
2.2.4 Klasifikasi Multilabel	33
2.2.5 BERT (Bidirectional Encoder Representation from Transformers)	34
2.2.6 Fitur Linguistik.....	37
2.2.7 SpaCy	38
2.2.8 POS (<i>Part of-Speech</i>) Tagging.....	39
2.2.9 NER (Named Entity Recognition).....	40
2.2.10 LSTM (Long-Short Term Memory).....	40
2.2.11 Bi-LSTM (Bidirectional Long Short-Term Memory).....	43

2.2.12	<i>Loss function</i> BCE (Binnary Cross Entrophy)	44
2.2.13	Evaluasi Model	46
2.2.14	Validasi Ahli	48
2.2.15	Fleiss' Kappa	48
2.2.16	Majority voting	50
2.2.17	Jaccard Similarity	50
2.2	Al-Qur'an	51
2.3	Integrasi Keilmuan	52
BAB III		57
3.1	Alur Penelitian (flowchart)	57
3.2	Tahapan Penelitian	57
3.2.1	Identifikasi Masalah	57
3.2.2	Pengumpulan Data	58
3.2.3	Splitting Dataset	60
3.2.4	Preprocessing	61
3.2.5	BERT <i>Embedding</i>	63
3.2.6	Fitur ekstraksi	65
3.2.7	Modelling	70
3.2.8	Evaluasi Model	74
3.3	Alat Penelitian	76
BAB IV		78
4.1	Pengumpulan Data	78
4.2	Preprocessing	80
4.2.1	Text Cleaning	80
4.2.2	Case Folding	81
4.2.3	Tokenization	82
4.3	BERT <i>Embedding</i>	83
4.4	Fitur Ekstraksi	86
4.4.1	Ekstraksi Pos Tagging	86
4.4.2	Ekstraksi NER	89
4.5	<i>Splitting</i> Dataset	91
4.6	Hasil Modelling Skenario I	93
4.7	Hasil Modelling Skenario II	95
4.8	Hasil Modelling Skenario III	97
4.9	Analisis Hasil Pemodelan	100
4.10	Analisis Eksperimen Skenario	105

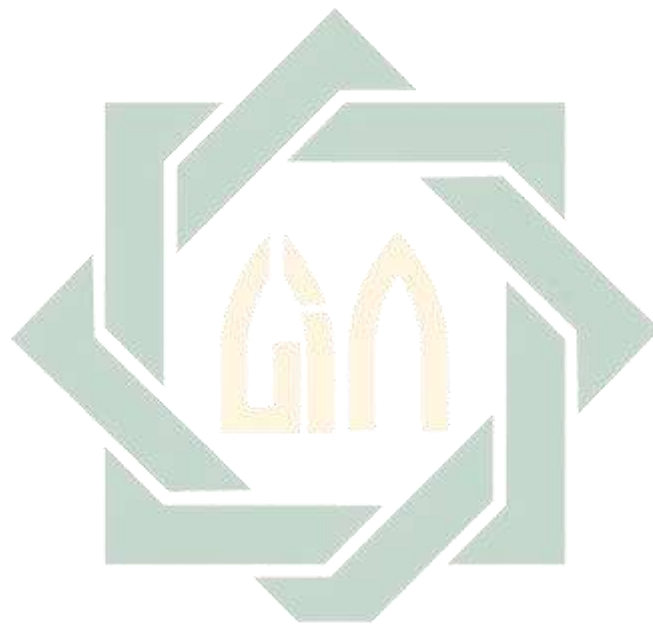
4.11	Analisis Imbalance Data	108
4.12	Analisis Validasi Pakar	110
4.12.1	Perhitungan Inter-Rate Reability	111
4.12.2	Majority voting	117
4.12.3	Perhitungan Jaccard Similarity	118
BAB V		122
5.1	Kesimpulan	122
5.2	Saran.....	122
Daftar Pustaka		124
Daftar Lampiran		132



UIN SUNAN AMPEL
S U R A B A Y A

DAFTAR GAMBAR

Gambar 2. 1 Fungsi Aktivasi Sigmoid (Akbar dkk, 2022)	30
Gambar 2. 2 Fungsi Aktivasi Tanh (Akbar dkk, 2022).....	31
Gambar 2. 3 Arsitektur BERT (Devlin dkk, 2019).....	35
Gambar 2. 4 <i>Embedding</i> Layer BERT (Devlin dkk, 2019).....	37
Gambar 2. 5 Arsitektur SpaCy (SpaCy, 2019).....	39
Gambar 2. 6 Arsitektur LSTM (Le dkk, 2019)	41
Gambar 2. 7 Arsitektur Bi-LSTM (Puteri, 2023).....	43
Gambar 3. 1 Tahapan Penelitian	57
Gambar 3. 2 Alur Pemodelan Sistem.....	70
Gambar 3. 3 Evaluasi Sistem	74
Gambar 4. 1 Distribusi Jumlah Data	78
Gambar 4. 2 Distribusi Multilabel dan Singlelabel.....	79
Gambar 4. 3 Distribusi <i>Embedding</i>	85
Gambar 4. 4 Distribusi Jumlah Tagging	88
Gambar 4. 5 Distribusi NER	90
Gambar 4. 6 Distribusi Splitting Dataset	92
Gambar 4. 7 Distribusi Split Per-Label.....	92
Gambar 4. 8 Grafik <i>Training Loss</i> Skenario-1	94
Gambar 4. 9 Grafik Validasi Hamming Loss.....	94
Gambar 4. 10 Grafik F1-Score.....	95
Gambar 4. 11 Grafik <i>Training Loss</i> Skenario-II.....	96
Gambar 4. 12 Grafik Validasi Hamming Loss	96
Gambar 4. 13 Grafik F1-Score.....	97
Gambar 4. 14 Grafik <i>Training Loss</i>	98
Gambar 4. 15 Grafik Validasi Hamming Loss	99
Gambar 4. 16 Grafik F1-Score.....	99
Gambar 4. 17 Grafik <i>Training Loss</i> 3 Skenario.....	100
Gambar 4. 18 Grafik <i>Training Validation Hamming Loss</i> 3 Skenario.....	101
Gambar 4. 19 Grafik Validasi Macro-F1 3 Skenario	101
Gambar 4. 20 Grafik Perbandingan 3 Skenario	102
Gambar 4. 21 Distribusi Kesepakatan per-item	113



UIN SUNAN AMPEL
S U R A B A Y A

DAFTAR TABEL

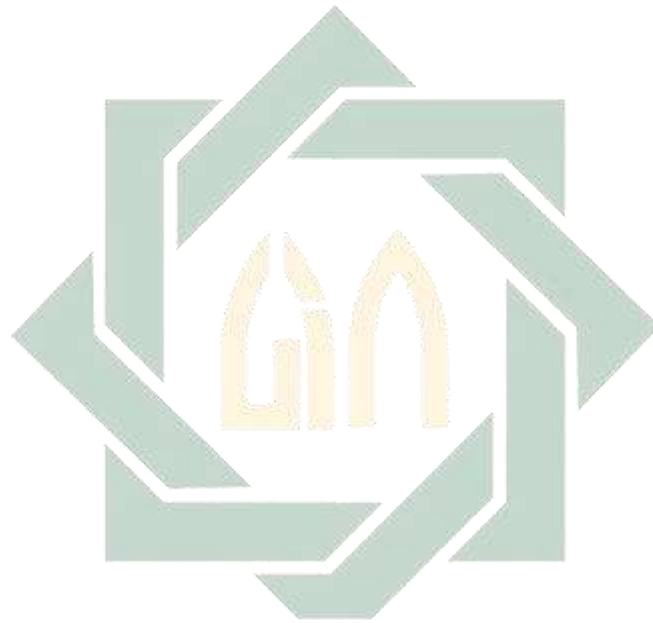
Tabel 2. 1 Penelitian Terdahulu	25
Tabel 3. 1 Kumpulan Dataset.....	58
Tabel 3. 2 Label Klasifikasi Topik Al-Qur'an.....	59
Tabel 3. 3 Ilustrasi Output BERT tokenizer.....	62
Tabel 3. 4 Ekstraksi BERT.....	63
Tabel 3. 5 Contoh <i>Embedding</i> dalam Dataset.....	64
Tabel 3. 6 Langkah BERT <i>Embedding</i>	64
Tabel 3. 7 Definisi POS-Taging.....	66
Tabel 3. 8 Ekstraksi POS Tagging.....	66
Tabel 3. 9 Contoh Ilustrasi Pos Tag.....	67
Tabel 3. 10 Definisi Entitas NER.....	68
Tabel 3. 11 Ekstraksi NER.....	68
Tabel 3. 12 Contoh Output NER.....	69
Tabel 3. 13 Skenario Penelitian	71
Tabel 3. 14 Hasil pos_weight setiap label.....	73
Tabel 3. 15 Alat Penelitian.....	77
Tabel 4. 1 Distribusi Dataset.....	80
Tabel 4. 2 Proses Cleaning.....	81
Tabel 4. 3 Proses Case Folding.....	81
Tabel 4. 4 Tahap Tokenisasi BERT	82
Tabel 4. 5 Parameter BERT Tokenisasi.....	83
Tabel 4. 6 Konfigurasi Tokenisasi	84
Tabel 4. 7 Distribusi Shape <i>Embedding</i> per Subset	84
Tabel 4. 8 Contoh <i>Embedding</i> BERT	85
Tabel 4. 9 Konfigurasi POS Tagging.....	86
Tabel 4. 10 Contoh Ekstraksi fitur POS Tagging	89
Tabel 4. 11 Konfigurasi NER.....	89
Tabel 4. 12 Contoh ekstraksi fitur NER.....	91
Tabel 4. 13 Performa <i>Training</i> dan <i>Testing</i>	93
Tabel 4. 14 Performa <i>Training</i> dan <i>Testing</i> Skenario II	95
Tabel 4. 15 Hasil Performa Model <i>Training</i> dan <i>Testing</i>	98

Tabel 4. 16 Tabel Perbandingan 3 Skenario	102
Tabel 4. 17 Tabel Perbandingan Penelitian Terdahulu	104
Tabel 4. 18 Tabel Analisis Penambahan POS Tagging	105
Tabel 4. 19 Tabel Analisis Penambahan NER	106
Tabel 4. 20 Tabel Matrik Setiap Skenario	107
Tabel 4. 21 Classification Report per Label Skenario III	108
Tabel 4. 22 Tabel Kategori Penilaian.....	111
Tabel 4. 23 Distribusi Kategori tiap Validator.....	112
Tabel 4. 24 Matriks Item.....	112
Tabel 4. 25 Perhitungan Proporsi tiap Kategori.....	114
Tabel 4. 26 Skala Interpretasi Fleiss' Kappa (Landis & Koch, 1977).....	115
Tabel 4. 27 Ringkasan Hasil Fleiss' Kappa.....	115
Tabel 4. 28 Kesepakatan Pairwise Antar-Validator.....	116
Tabel 4. 29 Pola Kesepakatan Voting Antar Validator.....	117
Tabel 4. 30 Hasil Perhitungan Jaccard.....	120

UIN SUNAN AMPEL
S U R A B A Y A

DAFTAR LAMPIRAN

Lampiran 1. Lembar Kesediaan Validator I.....	132
Lampiran 2. Pendukung Bidang Keahlian Validator I.....	133
Lampiran 3. Lembar Kesediaan Validator II	134
Lampiran 4. Pendukung Bidang Keahlian Validator II	135
Lampiran 5. Lembar Kesediaan Validator III	136
Lampiran 6. Pendukung Bidang Keahlian Validator III	137



UIN SUNAN AMPEL
S U R A B A Y A

Daftar Pustaka

- Abbot, D. (2013). *Introduction to Text mining*.
- Adeleke, A., Samsudin, N., Mustapha, A., & Ahmad Khalid, S. K. (2019). Automating quranic verses labeling using machine learning approach. *Indonesian Journal of Electrical Engineering and Computer Science*, 16, 925. <https://doi.org/10.11591/ijeecs.v16.i2.pp925-931>
- Aggarwal, C. C. (2018). *Neural networks and Deep learning*.
- Aiman, A., Lhaksana, K., & Jondri, J. (2022). Topic Classification of Quranic Verses in English Translation Using Word Centrality Measurement. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 6, 803–809. <https://doi.org/10.29207/resti.v6i5.4358>
- Alamro, H., Gojobori, T., Essack, M., & Gao, X. (2024). BioBBC: a multi-feature model that enhances the detection of biomedical entities. *Scientific Reports*, 14(1), 7697. <https://doi.org/10.1038/s41598-024-58334-x>
- Al-Omari, H., Abdullah, M. A., & Shaikh, S. (2020). *EmoDet2: Emotion Detection in English Textual Dialogue using BERT and BiLSTM Models*.
- Alsharef, A., Aggarwal, K., Garg, S., Koundal, D., Alyami, H., & Ameyed, D. (2022). An Automated Toxicity Classification on Social Media Using LSTM and Word Embedding. *Computational Intelligence and Neuroscience*, 2022. <https://doi.org/10.1155/2022/8467349>
- Ananda Pane, R., & Syahrul Mubarak, M. (2018). *KLASIFIKASI MULTI-LABEL PADA TOPIK AYAT AL-QURAN TERJEMAHAN BAHASA INGGRIS MENGGUNAKAN MULTINOMIAL NAIVE BAYES*.
- Anshori. (2013). *Ulumul Quran* (18 ed.). Rajawali Press.
- Artstein, R., & Poesio, M. (2008). Survey Article: Inter-Coder Agreement for Computational Linguistics. *Computational Linguistics*, 34(4), 555–596. <https://doi.org/10.1162/coli.07-034-R2>
- Atika. (2021). *Ekstraksi Informasi Berita Online dengan Named Entity Recognition (NER) dan Rule-Based untuk Visualisasi Penyakit Tropis di Indonesia*.
- Ayu Angraini, D., Ulya, A., Noviyanti, S., Jambi, U., Jambi -M uara Bulian NoKM, J., endalo Darat, M., Jambi Luar Kota, K., & uaro Jambi, K. M. (2024).

- Struktur Linguistik Bahasa Sintaksis Dan Sematik. *Journal on Education*, 07(01), 6664–6675.
- Bello, A., Ng, S. C., & Leung, M. F. (2023). A BERT *Framework* to Sentiment Analysis of Tweets. *Sensors*, 23(1). <https://doi.org/10.3390/s23010506>
- Brill, E. (2002). *A Simple Rule-Based Part of Speech Tagger*.
- Choi, S.-S., Cha, S.-H., & Tappert, C. C. (2010). *A Survey of Binary Similarity and Distance Measures*.
- Choirulfikri, M., Lhaksana, K., & Faraby, S. (2022). A Multi-Label Classification of Al-Quran Verses Using Ensemble Method and Naïve Bayes. *Building of Informatics, Technology and Science (BITS)*, 3, 473–479. <https://doi.org/10.47065/bits.v3i4.1287>
- Condorcet, M. J. A. N. (1785). *ProbabiliteDecisions*. Newwood Ltd.: International Association of Engineers.
- Cui, X. (2025). xiacui at SemEval-2025 Task 11: Addressing Data Imbalance in Transformer-Based Multi-Label Emotion Detection with Weighted Loss. Dalam S. Rosenthal, A. Rosá, D. Ghosh, & M. Zampieri (Ed.), *Proceedings of the 19th International Workshop on Semantic Evaluation (SemEval-2025)* (hlm. 247–256). Association for Computational Linguistics. <https://aclanthology.org/2025.semeval-1.37/>
- Devlin, J., Chang, M.-W., Lee, K., Google, K. T., & Language, A. I. (2019a). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. <https://github.com/tensorflow/tensor2tensor>
- Devlin, J., Chang, M.-W., Lee, K., Google, K. T., & Language, A. I. (2019b). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. <https://github.com/tensorflow/tensor2tensor>
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. <http://arxiv.org/abs/1810.04805>
- Ertugrul, A. M., & Karagoz, P. (2018). Movie Genre Classification from Plot Summaries Using *Bidirectional LSTM*. *2018 IEEE 12th International Conference on Semantic Computing (ICSC)*, 248–251. <https://doi.org/10.1109/ICSC.2018.00043>

- Fatiara, N., Safaat, N. H., Agustian, S., & Afrianty, I. (2024). *Komparasi Metode K-Nearest Neighbors Dan Long Short Term Memory* (Vol. 6, Nomor 2).
- Fleiss, J. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin*, 76, 378–382. <https://doi.org/10.1037/h0031619>
- Fleiss, J. L., Levin, B., & Cho Paik, M. (2003). *Statistical Methods for Rates and Proportions Third Edition*.
- Ghosh, A., Sufian, A., Sultana, F., Chakrabarti, A., & De, D. (2020). Fundamental Concepts of Convolutional Neural network. Dalam V. E. Balas, R. Kumar, & R. Srivastava (Ed.), *Recent Trends and Advances in Artificial Intelligence and Internet of Things* (hlm. 519–567). Springer International Publishing. https://doi.org/10.1007/978-3-030-32644-9_36
- Graves, A., & Schmidhuber, J. (2005). Framewise phoneme classification with bidirectional LSTM networks. *Proceedings. 2005 IEEE International Joint Conference on Neural networks, 2005.*, 4, 2047–2052 vol. 4. <https://doi.org/10.1109/IJCNN.2005.1556215>
- Grikopedia. (2026). *Named-entity recognition*. Grikopedia. https://grikopedia.com/page/Named-entity_recognition
- Grishman, R., & Sundheim, B. (1996). *Message Understanding Conference-6: A Brief History*.
- Gwet, K. L. (2014). *Handbook of Inter-rater reliability Fith Edition The Definitive Guide to Measuring the Extent of Agreement Among Raters Volume 1 Analysis of Categorical Ratings*. https://sites.fastspring.com/agreestat/instant/cac5ed978_1_7923_5463_2https://agreestat.com/books/cac5/https://agreestat.com/books/
- Hakim, M. F. (2014). *Membumikan Al-Qur'an*. Lirboy Press.
- Han, J. and, Kamber, M., & Pei, J. (2011). *Data Mining: Concepts and Techniques* (3 ed.).
- Hinojosa Lee, M. C., Braet, J., & Springael, J. (2024). Performance Metrics for Multilabel Emotion Classification: Comparing Micro, Macro, and Weighted F1-Scores. *Applied Sciences (Switzerland)*, 14(21). <https://doi.org/10.3390/app14219863>

- Hinton, G. E., Osindero, S., & Teh, Y.-W. (t.t.). *A fast learning algorithm for deep belief nets* *.
- Hinton, G., Srivastava, N., Krizhevsky, A., Sutskever, I., Rachmad, Y., & Salakhutdinov, R. (2014). *Dropout: A Simple Way to Prevent Neural networks from Overfitting. Journal of Machine learning Research, 15*, 1929–1958.
- Huda, N. S. (2019). *A Multi-label Classification on Topics of Quranic Verses (English Translation) Using Backpropagation Neural network with Stochastic Gradient Descent and Adam Optimizer*. IEEE.
- Huda, N. S., Mubarak, mohammad syahrul, & Adiwijaya. (2019). *A Multi-label Classification on Topics of Quranic Verses (English Translation) Using Backpropagation Neural network with Stochastic Gradient Descent and Adam Optimizer*. IEEE.
- Irfan, D., Gunawan, T., & Wanayumini, W. (2023). Comparison Of SGD, Rmsprop, and Adam Optimation In Animal Classification Using CNNs. *International Conference on Information Science and Technology Innovation (ICoSTEC)*, 2, 45–51. <https://doi.org/10.35842/icostec.v2i1.35>
- Irfan, R. A., & Lhaksana, K. M. (2023a). “Klasifikasi Ayat Al-Quran Terjemahan Bahasa Inggris Menggunakan Long Short Term Memory dan Bidirectional Long Short Term Memory.”
- Irfan, R. A., & Lhaksana, K. M. (2023b). *Klasifikasi Ayat Al-Quran Terjemahan Bahasa Inggris Menggunakan Long Short Term Memory dan Bidirectional Long Short Term Memory*.
- Isnain, A. R., Sihabuddin, A., & Suyanto, Y. (2020). *Bidirectional Long Short Term Memory Method and Word2Vec Extraction Approach for Hate Speech Detection. IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, 14(2), 169. <https://doi.org/10.22146/ijccs.51743>
- Izzaty, A. M. K., Mubarak, mohammad syahrul, Huda, N. S., & Adiwijaya. (2018). *A Multi-label Classification on Topics of Quranic Verses in English Translation Using Tree Augmented Naïve Bayes*. IEEE.
- Kholkovskaia, O. (2017). *Role of the Brown Corpus in the History of Corpus Linguistics*. <http://www.browncorpusmagazine.com/content/view/25>

- Kuswoyo. (2021). *Pengantar Studi Ilmu-ilmu Al-Qur'an*. PT. Nasya Expanding Managemen.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159—174. <https://doi.org/10.2307/2529310>
- Mardianti, S., Zidny Naf'an, M., & Hidayatulloh, I. (2018). *EKSTRAKSI TF-IDF N-GRAM DARI KOMENTAR PELANGGAN PRODUK SMARTPHONE PADA WEBSITE E-COMMERCE*.
- Ma'rifah, H., Prasetya Wibawa, A., & Akbar, M. I. (2020). *Sains, Aplikasi, Komputasi dan Teknologi Informasi Klasifikasi artikel ilmiah dengan berbagai Skenario preprocessing*. 2(2), 70.
- Maryanto. (2026). *Mengenal Natural Language Processing (NLP): Konsep dasar, komputasi linguistik, dan tantangannya*. Medium. <https://medium.com/@sitisalsamar.03/mengenal-natural-language-processing-nlp-konsep-dasar-komputasi-linguistik-dan-tantangannya-57c845d1b299>
- Mowafy, M., Rezk, A., & El-Bakry, H. (2018). An Efficient Classification Model for Unstructured Text Document. *American Journal of Computer Science and Information Technology*, 06. <https://doi.org/10.21767/2349-3917.100016>
- Mubarok, mohammad syahrul, Izaaty, A. mira K., & Huda, N. S. (2018). *A Multi-label Classification on Topics of Quranic Verses in English Translation Using Tree Augmented Naïve Bayes*. IEEE.
- Mutasa, S., Sun, S., & Ha, R. (2020). Understanding artificial intelligence based radiology studies: What is *overfitting*? *Clinical Imaging*, 65, 96–99. <https://doi.org/10.1016/j.clinimag.2020.04.025>
- Mutmainnah, S., Ansyor Lorosae, T., & Ramadhan, S. (2025). *Model Text Embedding dan TF-IDF+Ngram untuk Meningkatkan Kinerja Algoritma Binary Classifier pada Klasifikasi SMS Palsu*. 4(1), 55–64. <https://ojs.trigunadharma.ac.id/index.php/jsi>
- Nadeau, D., & Sekine, S. (2007a). A survey of *Named Entity Recognition* and classification. *Lingvisticæ Investigationes*, 30(1), 3–26. <https://doi.org/https://doi.org/10.1075/li.30.1.03nad>

- Nadeau, D., & Sekine, S. (2007b). *A survey of Named Entity Recognition and classification*. <http://projects.ldc.upenn.edu/gale/>
- Nurfajiah, A., Hartati, T., Amalia, D. R., & Cirebon, S. I. (2021). *Integrated Library System untuk Meningkatkan Efektivitas Layanan Perpustakaan Dengan Menggunakan Metode Algoritma Apriori (Studi Kasus: Perpustakaan Kabupaten Cirebon)*.
- Onan, A. (2023). Hierarchical graph-based text classification framework with contextual node embedding and BERT-based dynamic fusion. *Journal of King Saud University - Computer and Information Sciences*, 35(7), 101610. <https://doi.org/https://doi.org/10.1016/j.jksuci.2023.101610>
- Pratama, E. D. (2022). Implementasi Model Long-Short Term Memory (LSTM) pada Klasifikasi Teks Data SMS Spam Berbahasa Indonesia. *The Journal on Machine learning and Computing Intelligence (JMLCI)*, 1(2), 38–42.
- Purbo, O. W. (2019). *Text mining, Analisis Medsos, Kekuatan Brand dan Intelijen di Internet*.
- Qamar, R., & Zardari, B. A. (2023). *Artificial neural networks: An Overview. Mesopotamian Journal of Computer Science*, 2023, 124–133. <https://doi.org/10.58496/MJCSC/2023/015>
- Ramachandran, P., Zoph, B., & Le, Q. V. (2017). *Searching for Activation functions*. <http://arxiv.org/abs/1710.05941>
- Ramadhani, A., Safaat, N., Agustian, S., Iskandar, I., & Sanjaya, S. (2025). Perbandingan Performa Metode Klasifikasi Teks Multilabel Hadis Terjemahan Bukhari Menggunakan Support Vector Machine dan Long Short Term Memory. *MALCOM: Indonesian Journal of Machine learning and Computer Science*, 5(3), 896–907. <https://doi.org/10.57152/malcom.v5i3.2051>
- Rofiqul Muslikh, A., Akbar, I., Rosal Ignatius Moses Setiadi, D., & Md Mehedul Islam, H. (2024). Multi-label Classification of Indonesian Al-Quran Translation based CNN, BiLSTM, and FastText. Dalam *Februari* (Vol. 23, Nomor 1). <https://quran.kemenag.go.id>.
- Ruang Jurnal. (2025). *Memahami validasi expert judgement: Konsep, proses, dan implikasinya*. Ruang Jurnal.

- Sagala, L. O. S., & Sagala, A. (2022). *Klasifikasi Cats dan Dogs*.
- Sari, W. K., Rini, D. P., & Malik, R. F. (2020). Text Classification Using Long Short-Term Memory With GloVe Features. *Jurnal Ilmiah Teknik Elektro Komputer dan Informatika*, 5(2), 85. <https://doi.org/10.26555/jiteki.v5i2.15021>
- Sari, Wierfi, A. D., & Setyanto, A. (2019). Sentiment Analysis of Customer Satisfaction on Transportation Network Company Using Naive Bayes Classifier. *2019 International Conference on Computer Engineering, Network, and Intelligent Multimedia (CENIM)*, 1–6. <https://doi.org/10.1109/CENIM48368.2019.8973262>
- Scikit Learn. (2024). *Scikit-learn documentation – Classification metrics*.
- Sebastiani, F. (2001). *Machine learning in Automated Text Categorization*. <http://iinwww.ira.uka.de/bibliography/Ai/automated.text.categorization.html>
- Sheng, V. S. , Provost, F. , & Ipeirotis, P. G. (2008). *Get another label? Improving data quality and data mining using multiple, noisy labelers*. ACM.
- SpaCy. (2022). *Spacy* . <https://spacy.io/usage/spacy-101>. <https://spacy.io/usage/spacy-101>
- Sugiyono. (2016). *Metode penelitian kuantitatif, kualitatif, dan R&D*. Alfabeta.
- Sumathy, K., & Chidambaram, M. (2013). Text mining: Concepts, Applications, Tools and Issues An Overview. *International Journal of Computer Applications*, 80, 29–32. <https://doi.org/10.5120/13851-1685>
- Sun, R. (2019). *Optimization for deep learning: theory and algorithms*. <http://arxiv.org/abs/1912.08957>
- Tri Hermanto, D., Setyanto, A., & Luthfi, E. T. (2021). *Algoritma LSTM-CNN untuk Sentimen Klasifikasi dengan Word2Vec pada Media Online LSTM-CNN Algorithm for Sentiment Clasifcation with Word2Vec On Online Media*.
- Trisiawan, I. K., & Yuliza, Y. (2022). Penerapan Multi-Label Image Classification Menggunakan Metode Convolutional Neural network (CNN) Untuk Sortir Botol Minuman. *Jurnal Teknologi Elektro*, 13(1), 48. <https://doi.org/10.22441/jte.2022.v13i1.009>

- Wu, Y., & Liu, L. (2023). Selecting and Composing *Learning rate* Policies for Deep Neural networks. *ACM Transactions on Intelligent Systems and Technology*, 14(2). <https://doi.org/10.1145/3570508>
- Wu, Y., Liu, L., Bae, J., Chow, K.-H., Iyengar, A., Pu, C., Wei, W., Yu, L., & Zhang, Q. (2019). *Demystifying Learning rate Policies for High Accuracy Training of Deep Neural networks*. <http://arxiv.org/abs/1908.06477>
- Zahra, E., Sibaroni, Y., & Prasetyowati, S. (2023). Classification of Multi-Label of Hate Speech on Twitter Indonesia using LSTM and BiLSTM Method. *JINAV: Journal of Information and Visualization*, 4, 170–178. <https://doi.org/10.35877/454RL.jinav1864>
- Zhang, D., & Mo, J. (2025). *LinguaSynth: Heterogeneous Linguistic Signals for News Classification*. <http://arxiv.org/abs/2506.21848>
- Zhu, Z., Zhang, M., & Jiang, X. (2021). *Fingerspelling Identification for Chinese Sign Language via Wavelet Entropy and Kernel Support Vector Machine* (hlm. 539–549). https://doi.org/10.1007/978-981-15-5679-1_52

UIN SUNAN AMPEL
S U R A B A Y A