

PENDEKATAN *BIDIRECTIONAL* LSTM UNTUK KLASIFIKASI FUNGSIONAL DAN NON-FUNGSIONAL *REQUIREMENTS* BERBASIS *NATURAL LANGUAGE PROCESSING*

Rizki Rizkyatul Basir¹, Yunita Endra Megiati², Noor Komari Pratiwi³

^{1,2,3}Universitas Indraprasta PGRI

Jalan Nangka Raya, C Jl. TB Simatupang No.58, Tj. Bar., Kec. Jagakarsa, Jakarta Selatan
rizkiyatulbasir@gmail.com, yunitaendra@gmail.com, noorkomaripratiwi01@gmail.com

ABSTRAK

Proses klasifikasi *Functional Requirement* (FR) dan *Non-Functional Requirement* (NFR) merupakan tahap penting dalam *Software Engineering* untuk memastikan sistem yang dibangun memenuhi aspek fungsional maupun kualitas. Tantangannya terletak pada proses manual, sehingga menyita waktu dan rentan terhadap *human error*. Penelitian ini bertujuan untuk mengotomatiskan proses klasifikasi dengan menggunakan pendekatan *Natural Language Processing* (NLP) berbasis model *Bidirectional Long Short-Term Memory* (BiLSTM). Metode ini dipilih karena mampu memahami konteks kalimat dari dua arah, sehingga meningkatkan akurasi dalam pengenalan teks *requirement*. Dataset yang digunakan diambil dari repositori Mendeley Data dengan jumlah 6.117 entri, kemudian melalui tahap *pre-processing* menjadi 6.086 data valid dengan 3.964 FR dan 2.122 NFR. Data diproses dengan tahapan *tokenisasi*, *padding*, *encoding label*, serta pembagian data latih dan uji. Model BiLSTM dilatih menggunakan TensorFlow atau Keras dan dibandingkan dengan pendekatan TF-IDF + Logistic Regression sebagai *baseline*. Hasil uji menunjukkan model BiLSTM mencapai akurasi 84,31%, dengan performa yang konsisten dalam mengenali kelas *requirement*. Penelitian ini menunjukkan efektivitas pendekatan *deep learning* efektif untuk membantu proses *requirement engineering* untuk meningkatkan efisiensi sekaligus meminimalkan *human error*.

Kata Kunci: Requirements Engineering, FR, NFR, NLP, BiLSTM

ABSTRACT

The process of classifying *Functional Requirements* (FR) and *Non-Functional Requirements* (NFR) is an important stage in *Software Engineering*, ensuring that the system being built meets both functional and quality requirements. The challenge lies in the manual process, which is time-consuming and prone to human error. This study aims to automate the classification process using a *Natural Language Processing* (NLP) approach based on the *Bidirectional Long Short-Term Memory* (BiLSTM) model. This method was chosen because it can understand the context of a sentence from two directions, thereby improving the accuracy of requirement text recognition. The dataset was obtained from the Mendeley Data repository, comprising 6,117 entries, and underwent preprocessing, yielding 6,086 valid data points: 3,964 FR and 2,122 NFR. The data was processed through tokenization, padding, label encoding, and training and testing data division stages. The BiLSTM model was trained with TensorFlow or Keras and compared with the TF-IDF + Logistic Regression baseline. The test results showed that the BiLSTM model achieved an accuracy of 84.31%, with consistent performance in recognizing the requirement class. This study demonstrates the effectiveness of the deep learning approach in supporting the requirements engineering process, improving efficiency while minimizing human error.

Keyword: Requirements Engineering, FR, NFR, NLP, BiLSTM

PENDAHULUAN

Dalam proses pengembangan perangkat lunak, kebutuhan sistem (system requirements) terbagi menjadi dua kategori utama yaitu *Functional Requirement* (FR) dan *Non-Functional Requirement* (NFR). FR menjelaskan fungsi yang harus dimiliki sistem, sedangkan NFR mendeskripsikan karakteristik kualitas seperti performa, kehandalan, keamanan, dan kegunaan (Casamayor et al., 2010). Pemisahan keduanya sangat penting untuk menjamin

sistem memenuhi aspek fungsional sekaligus kualitas.

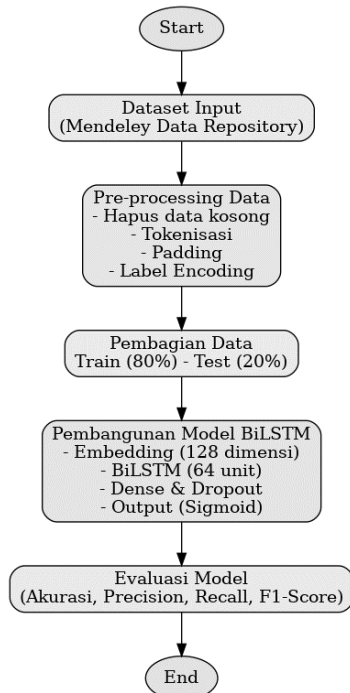
Namun, proses klasifikasi requirement umumnya dilakukan secara manual oleh seorang *system analys*. Kemudian Hal ini menimbulkan permasalahan seperti waktu pengerjaan yang panjang, subjektivitas, serta inkonsistensi antar analis. Seiring perkembangan teknologi, pendekatan berbasis *Natural Language Processing* (NLP) dan deep learning dapat dimanfaatkan untuk

mengotomatisasi klasifikasi requirement secara lebih cepat dan konsisten. Penelitian ini mengusulkan penerapan model *Bidirectional Long Short-Term Memory* (BiLSTM) dalam klasifikasi FR dan NFR. Model deep learning, termasuk BiLSTM dan varian berbasis pre-trained language models, telah terbukti meningkatkan akurasi klasifikasi *requirement* (Abbas et al., 2024; Song et al., 2022; Vijayvargiya et al., 2022).

METODE PENELITIAN

Penelitian dilakukan dengan menggunakan bahasa pemrograman Python dengan Jupyter Notebook sebagai *Development Environment*. Beberapa pustaka yang digunakan meliputi *pandas* untuk pengolahan data, *scikit-learn* untuk encoding label dan evaluasi, *TensorFlow* atau *Keras* untuk pembangunan model BiLSTM dan yang terakhir *matplotlib* dan *seaborn* untuk visualisasi hasil.

Berikut merupakan tahapan yang dilakukan dalam melakukan penelitian dalam pengujian:



Gambar 1. Flowchart Pengujian

- Dataset:** Untuk data yang digunakan diperoleh dari repositori Mendeley Data. Dataset ini berisi kumpulan kalimat kebutuhan perangkat lunak dalam bahasa Inggris dengan label FR atau NFR. Jumlah entri awal sebanyak 6.117 baris

lalu setelah pembersihan data didapat 6.086 baris valid dengan kategori FR = 3.964 dan NFR = 2.122. Contoh requirement:

Tabel 1. Gambar Tabel Requirement

=== Tabel Requirement: FR ===

	Requirement Text	Type
241	User shall be able to reset their password by ...	FR
1571	Users shall add specific stocks to their portf...	FR
1968	System shall provide a search facility for adm...	FR
1153	Program Administrators and Nursing Staff Membe...	FR
434	The system shall be able to perform basic and ...	FR

=== Tabel Requirement: NFR ===

	Requirement Text	Type
4516	Support solution shall be accessible both with...	NFR
1104	Staff members shall be able to complete a set ...	NFR
857	The administrator shall have the ability to ed...	NFR
5114	System shall optimize query execution for perf...	NFR
899	The administrator shall have the ability to bl...	NFR

Jumlah data yang digunakan:

$$N = \text{Jumlah dataset} \quad (1)$$

- Pre-processing:** Menghapus baris tanpa label, Tokenisasi teks, Padding/truncating hingga 100 token, Encoding label (FR=0, NFR=1). Rumus sederhana jumlah data bersih setelah pembersihan:

$$X' = N \times (m - k) \quad (2)$$

Dik: N = total entri awal, k = jumlah entri yang dihapus.

- Pembangunan Model:** Data dibagi menjadi 80% (4.868) data latih dan 20% (1.218) data uji secara stratified split. Proporsi pembagian:

$$\text{Data Latih} = N \times p \quad (3)$$

$$\text{Data Uji} = N \times (1 - p) \quad (4)$$

Dik: $p = 0.8$

Model utama menggunakan *Bidirectional LSTM* (BiLSTM) dengan arsitektur:

- Embedding layer (128 dimensi)
- BiLSTM (64 unit)
- Dense layer (64 neuron, ReLU)
- Dropout 0.5 dan 0.3
- Output layer (Sigmoid)

Sebagai baseline, digunakan pendekatan TF-IDF + Logistic Regression.

4. **Evaluasi Model:** Evaluasi dilakukan menggunakan metrik: akurasi, precision, recall, dan F1-score.

$$Akurasi = \frac{Prediksi\ Benar}{Total\ data} \quad (5)$$

$$Precision = \frac{TP}{TP+FP} \quad (6)$$

$$Recall = \frac{TP}{TP+FNTP} \quad (7)$$

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (8)$$

HASIL DAN PEMBAHASAN

Setelah dilakukan pelatihan model dengan arsitektur BiLSTM, diperoleh hasil uji pada data testing dengan akurasi sebesar 84,31%. Nilai ini menunjukkan bahwa model mampu mengklasifikasikan requirement ke dalam kategori *Functional Requirement* (FR) dan *Non-Functional Requirement* (NFR) dengan tingkat ketepatan yang cukup baik.

Berikut merupakan Informasi dataset:

```

=== Informasi Dataset ===
Total baris: 6117
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 6117 entries, 0 to 6116
Data columns (total 2 columns):
#   Column          Non-Null Count  Dtype
---  -
0   Requirement Text  6117 non-null   object
1   Type              6086 non-null   object
dtypes: object(2)
    
```

Gambar 2. Informasi Dataset

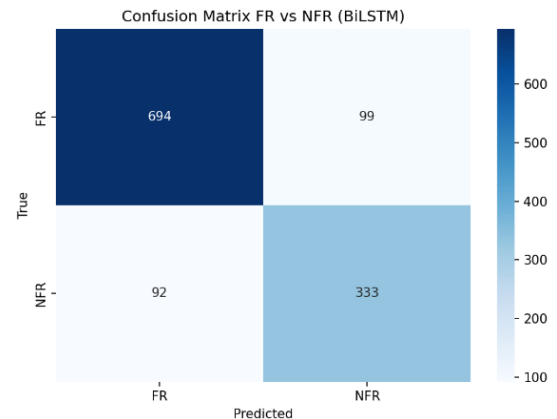
Pada tahap awal, distribusi data setelah pre-processing memperlihatkan bahwa jumlah FR lebih besar dibanding NFR (FR = 3.964 dan NFR = 2.122). Ketidakseimbangan ini berdampak pada hasil klasifikasi, di mana model lebih mudah mengenali FR dibandingkan NFR.

Tabel 2. Tabel Hasil Pre-processing

=== Distribusi Kelas ===	
Type	Jumlah
FR	3964
NFR	2122

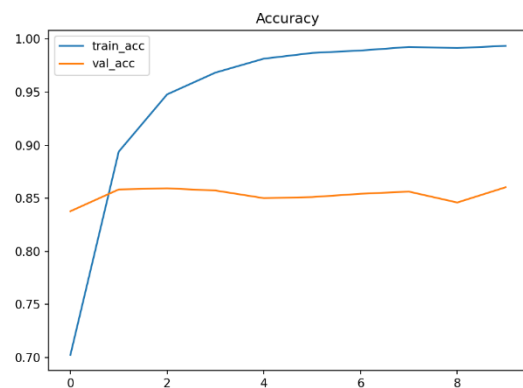
Evaluasi model menggunakan confusion matrix menunjukkan sebagian besar FR dapat

diprediksi dengan benar, sedangkan pada kelas NFR masih ditemukan beberapa kesalahan prediksi. Hal ini wajar mengingat jumlah data NFR relatif lebih sedikit.



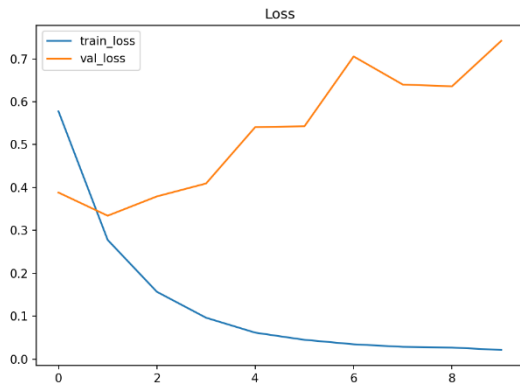
Gambar 3. Gambar Confision Matrix

Kurva akurasi memperlihatkan tren peningkatan pada data latih hingga melebihi 95%, sedangkan akurasi validasi stabil di kisaran 84%. Hal ini mengindikasikan bahwa model berhasil belajar pola dengan baik, meskipun terdapat perbedaan antara akurasi training dan validasi.



Gambar 4. Gambar Kurva Akurasi

Sementara itu, kurva loss menunjukkan bahwa nilai training loss terus menurun, sedangkan validation loss mulai meningkat setelah epoch ke-3. Pola ini mengindikasikan adanya gejala overfitting, meskipun performa validasi masih relatif stabil.



Gambar 5. Gambar Kurva Loss

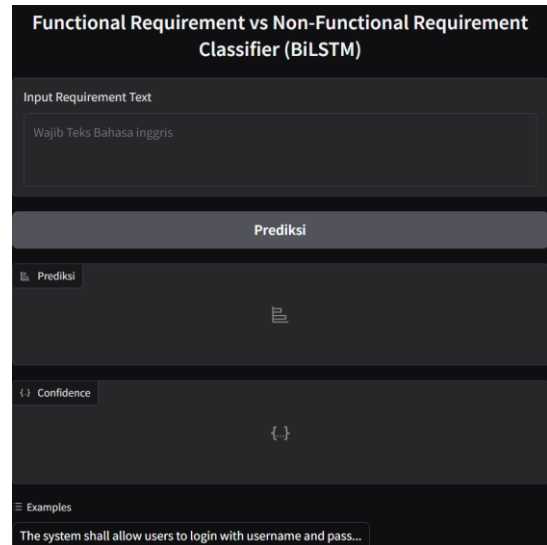
Dari sisi metrik evaluasi, nilai precision, recall, dan F1-score pada kelas FR lebih tinggi dibandingkan dengan NFR. Hal ini kembali dikaitkan dengan proporsi data yang tidak seimbang. Namun, secara keseluruhan, model BiLSTM tetap mampu memberikan hasil yang lebih baik dibanding baseline TF-IDF + Logistic Regression yang hanya mencapai akurasi sekitar 86%.

Tabel 3. Tabel Hasil TF-IDF+Logistic Reg

=== Classification Report: TF-IDF + Logistic Regression

	precision	recall	f1-score	support
FR	0.86	0.94	0.90	793
NFR	0.87	0.71	0.78	425
accuracy			0.86	1218
macro avg	0.86	0.83	0.84	1218
weighted avg	0.86	0.86	0.86	1218

Sebagai bukti penerapan, model juga diuji pada requirement menggunakan antarmuka interaktif berbasis Gradio. Hasil pengujian menunjukkan bahwa model mampu memberikan prediksi yang konsisten, misalnya requirement “The application must provide a search feature for products” diklasifikasikan sebagai FR dengan tingkat keyakinan yang tinggi, sedangkan requirement “The software shall comply with data privacy regulations” dikenali sebagai NFR.



Gambar 6. Gambar Gradio (Web Form)

Secara umum, hasil penelitian ini menegaskan bahwa penggunaan BiLSTM efektif untuk mengotomatisasi klasifikasi FR dan NFR. Model ini berpotensi digunakan sebagai tool pendukung dalam proses requirement engineering, sehingga pekerjaan analisis kebutuhan dapat dilakukan lebih cepat dan lebih konsisten.

SIMPULAN DAN SARAN

Penelitian ini membuktikan bahwa pendekatan BiLSTM dapat digunakan untuk mengklasifikasikan requirement ke dalam kategori FR dan NFR dengan akurasi 84,31%. Hasil akurasi menunjukkan performa yang kompetitif bila dibandingkan dengan penelitian terdahulu yang menggunakan machine learning berbasis fitur linguistic (Casamayor et al., 2010; Kurtanovic & Maalej, 2017), meskipun masih berada di bawah capaian model berbasis pre-trained language models (Song et al., 2022). Dibandingkan baseline TF-IDF + Logistic Regression 86,04%, model BiLSTM memberikan performa yang lebih seimbang terutama dalam mendeteksi NFR.

Saran terkait penelitian ini yaitu model dapat dikembangkan lebih lanjut dengan data augmentation, teknik regulasi untuk mengurangi *overfitting*, dan dukungan dataset multibahasa.

UCAPAN TERIMA KASIH

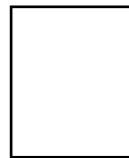
Penulis mengucapkan terima kasih kepada Universitas Indraprasta PGRI dan pihak-pihak yang telah memfasilitasi dan mendukung

penelitian ini dengan data dan fasilitas yang diperlukan.

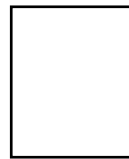
DAFTAR PUSTAKA

- Abbas, J., Hu, Z., Kanwal, S., Ahmad, A., Almogren, A., & Altameem, A. (2024). *Bi-LSTM-Based Model for Classifying Software Requirements*. <https://doi.org/10.20944/preprints202410.2129.v1>
- Casamayor, A., Godoy, D., & Campo, M. (2010). Identification of non-functional requirements in textual specifications: A semi-supervised learning approach. *Information and Software Technology*, 52(4), 436–445. <https://doi.org/10.1016/j.infsof.2009.10.010>
- Kurtanovic, Z., & Maalej, W. (2017). Automatically Classifying Functional and Non-functional Requirements Using Supervised Machine Learning. *Proceedings - 2017 IEEE 25th International Requirements Engineering Conference, RE 2017*, 490–495. <https://doi.org/10.1109/RE.2017.82>
- Song, D., Gao, S., He, B., & Schilder, F. (2022). On the Effectiveness of Pre-Trained Language Models for Legal Natural Language Processing: An Empirical Study. *IEEE Access*, 10, 75835–75858. <https://doi.org/10.1109/ACCESS.2022.3190408>
- Vijayvargiya, S., Kumar, L., Murthy, L. B., & Misra, S. (2022). Software Requirements Classification using Deep-learning Approach with Various Hidden Layers. *Proceedings of the 17th Conference on Computer Science and Intelligence Systems, FedCSIS 2022*, 895–904. <https://doi.org/10.15439/2022F140>
- Arora, C., Sabetzadeh, M., Briand, L. C., & Zimmer, F. (2019). *Functional and Non-functional Requirements Dataset*. Mendeley Data. <https://data.mendeley.com/datasets/4ysx9fyzv4/1>

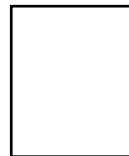
Biografi Penulis



Biografi Penulis 1 Rizki Rizkyatul Basir, Universitas Indraprasta PGRI, Magister Ilmu Komputer, Software Engineering atau Machine Learning.



Biografi Penulis 2 Yunita Endra Megiati, Universitas Indraprasta PGRI, Magister Pendidikan, kewarganegaraan dan perubahan sosial



Biografi Penulis 3 Noor Komari Pratiwi, Universitas Indraprasta PGRI, Magister Pendidikan, komunikasi dan literasi gizi