Comparison of News Text Summarization Using NLTK and TextRank Based on Python Programming

Artika Surniandari[1], Hilda Rachmi*[2], Ahmad Fauzi[3], Lila Dini Utami[3]

[1,2,3,4]Universitas Bina Sarana Informatika

ARTICLE INFORMATION	ABSTRACT
	Text summarization technology is increasingly used to simplify the vast amount of news information available in the digital era. This study compares two popular text summarization methods, the Natural Language Toolkit (NLTK) and TextRank, in the context of news summarization using the Python programming language. The goal of this research is to evaluate the performance of both algorithms based on summary quality and processing time. The dataset comprises a collection of news articles in Indonesian, processed using both methods. The results indicate that each algorithm has distinct advantages: TextRank excels in generating more coherent summaries, while NLTK demonstrates faster processing times. This study aims to contribute insights into the selection of an appropriate text summarization method for automating news summarization across various applications.

Corresponding Author:

Hilda Rachmi,
Sistem Informasi Kampus Kota Bogor,
Universitas Bina Sarana Informatika,
Jl. Kramat Raya no 98 Jakarta Pusat
Email: hilda.hlr@bsi.ac.id

INTRODUCTION

In the digital era, the rapid increase in news production and dissemination has led to an overwhelming amount of information that can be difficult for readers to fully absorb. Accessibility to the continuously expanding volume of online information would have a significant impact without summarization (Alami Merrouni et al., 2023). Automatic text summarization, a branch of natural language processing (NLP), has become an essential tool for condensing information, allowing users to quickly grasp the essence of news without needing to read the entire content. This method not only saves time but also enhances the accessibility and usability of information, especially in journalism, education, and business. Automatic text summarization is crucial in the era of unstructured big data, where the volume of textual data has increased significantly (Setiawan & Alexander, 2024).

Natural language processing (NLP) is a field of linguistics that enables machines to understand human language(Sadam et al., 2024). NLP processes complex, unstructured text data, including tasks such as word segmentation, part-of-speech tagging, named entity recognition, relation extraction, and machine translation. Representation learning helps represent natural language data at various levels and across multiple domains, thus enhancing the effectiveness and resilience of NLP performance. NLP uses theories and technologies to analyze

text content, enabling human-computer interaction through conversation and allowing computers to recognize, process, and respond to commands.

Among the various techniques available for text summarization, the Natural Language Toolkit (NLTK) and TextRank are two widely used approaches, each offering different methods for generating concise summaries. NLTK, a suite of symbolic and statistical NLP libraries and programs, utilizes extractive summarization techniques in which key sentences are directly selected from the source text. On the other hand, TextRank, an unsupervised algorithm inspired by Google's PageRank, models sentences as nodes in a graph and ranks them based on their relevance, identifying the most informative segments in the text. TextRank generates a fully connected, undirected graph of sentences or words in a document, where each sentence is represented by a node (Setiawan & Alexander, 2023).

This study aims to compare NLTK and TextRank summarization techniques by analyzing their effectiveness in summarizing news using Python programming. A system software in the form of a database can be used for data acquisition processes utilizing natural language processing (NLP) (Jimly Hanif et al., 2023). NLP is part of machine learning related to text learning (Mahawardana et al., 2022). NLP is used by Sentiment Analysis to determine the sentiment in a dataset. A large amount of data generated today is unstructured, which involves processing to generate information (Umunyana et al., 2024).

NLTK is one of the most commonly used Python libraries for natural language processing, allowing quick and easy processing of natural language text. NLTK has a large number of built-in corpora containing various types of textual material, such as novels, news, network chat texts, and movie reviews (Wang & Hu, 2021). NLTK and Keras are two powerful libraries frequently used in NLP (Kumar et al., 2024). Python is one of the programming languages commonly used by programmers to write programs with certain rules that must be met to predict errors and issues in the created program (Mahawardana et al., 2022).

NLTK is a specialized library that supports text stemming processing using the Python programming language (Albab et al., 2023). NLTK provides tools for preprocessing, such as tokenization and stopword removal, but usually does not include ranking methods like TextRank, while TextRank is a graph-based algorithm similar to PageRank that ranks sentences based on similarity, enabling the identification of the most important sentences in the context of news texts.

TextRank is a graph-based algorithm that ranks text by representing objects within the text. TextRank determines whether a sentence is important based on the structure of a graph. The first step is to create a graph consisting of points (nodes) representing each set (Pertiwi, 2022). TextRank is one of the string matching algorithms that simplifies and speeds up word searches (Munthe, 2022).

By examining these methods in terms of summary quality and processing efficiency, this article provides insights into the strengths and limitations of each approach, offering valuable information for choosing the most suitable text summarization technique in practical applications.

The objective of this study is to compare the results of NLTK and TextRank methods to highlight the differences in approach in: a. **Extractive summarization**: where both methods produce summaries by selecting sentences from the original text but use different approaches and algorithms. b. **Effectiveness on news texts**: by testing the methods directly on news texts, providing a direct perspective on how these methods work for information-dense texts such as news. c. **Practical Algorithm Comparison**: This article aims to outline the strengths and weaknesses of each method by evaluating the summary results practically:

Previous studies on text summarization have been widely conducted, especially in the context of news or long-text summarization. The following table presents some relevant findings from previous research:

Table 1. Previous Research on NLTK and TextRank

Title	Results
Implementation of Text Summarization on App Reviews Using the Maximum Marginal Relevance Method (Louis et al., 2022)	The research results using the MMR method produced an average precision score of 40.4% over three trials, with the highest precision value being 60.4% achieved in an experiment using a parameter value = 0.7.

Summarizing Indonesian Language Articles Using Textrank with BM25 Weighting (Hernawan et al., 2022)	In this study, each sentence in a document was ranked, and the top-ranked sentences were selected as the summary. The similarity function used was BM25 with the PageRank ranking method. The study achieved optimal summary quality with a compression rate of 30%, yielding average precision, recall, and f-measure scores of 0.552, 0.552, and 0.552, respectively.
Text Summarization on Soccer News Using Python-Based Natural Language Toolkit (NLTK) Library (Rifano et al., 2020)	This research aimed to assist readers in summarizing news, particularly soccer news, by using the NLTK library in Python programming. The results showed that out of 15 original sentences, the NLTK library summarized the text down to 6 core sentences.
Application of the TextRank Algorithm in Summarizing Word and PDF Text (Setiawan & Alexander, 2024)	This study discusses the application of the TextRank algorithm to summarize Word and PDF texts. The method used includes text preprocessing and the TextRank algorithm.
Variations of the Similarity Function of TextRank for Automated Summarization (Barrios et al., 2016)	*Barrios et al. (2016)*, demonstrated the effectiveness of TextRank in cross-linguistic summarization, especially for non-English texts. The results indicate that TextRank has strong adaptability as it focuses more on similarity patterns rather than mere word frequency. This method is flexible and suitable for application to news texts, which have varied language styles.

This study compares news text summarization using NLTK and TextRank in Python programming, employing a rarely used technique. This research provides a direct comparative analysis between NLTK and TextRank, offering deeper insights into the effectiveness and performance of each method within the news context. NLTK is used for preprocessing and TextRank for evaluation, exploring the potential for combining these methods to produce high-quality summaries that could serve as a reference for future hybrid technique development.

RESEARCH METHOD

In this study, the research methodology includes the following steps:

1. **Data Collection** This comparative study aims to evaluate the effectiveness of two summarization methods (NLTK and TextRank) in generating summaries of news text using a quantitative approach to measure the performance of each method based on measurable metrics. Data sources are collected from various sources, such as online news sites and publicly available news across multiple themes, to obtain diverse representations. Data can be stored in text or CSV format, with columns containing the news title, content, and category.

2. **Method Implementation**

   a. **Preprocessing with NLTK**
      Using NLTK to perform text preprocessing, including sentence and word tokenization, stopword removal, and stemming.
   b. **TextRank**
      Applying the TextRank algorithm to the same data, calculating sentence similarity to build a graph and compute sentence scores.

c. **Code Implementation**
Developing Python code for both methods, using required libraries (such as NLTK and networkx for TextRank).

3. **Testing and Evaluation** After obtaining summaries from both methods, evaluation metrics will be used to measure summary quality. This involves subjective assessment by users or experts to evaluate the relevance and quality of the summaries. Performance analysis is conducted by comparing the performance of the two methods based on predefined metrics.

4. **Data Analysis** Drawing conclusions based on the analysis performed, including the strengths and weaknesses of each method, and providing recommendations for further research or summarization-based application development.

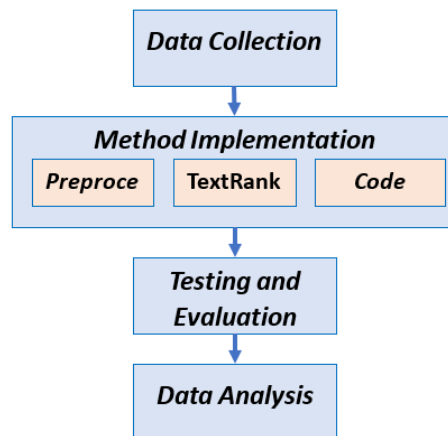The research steps are illustrated in Figure 1.

Figure 1. Research Stages

RESULTS AND DISCUSSION

NLTK is a widely used library in Python for natural language processing. In the summarization process, NLTK enables us to break down text into sentences, calculate word frequency, and select key sentences based on word frequency.

The general steps in summarization using NLTK are as follows:

1. **Word Frequency Calculation**: Calculating how often each word appears.

```python
# Frequency Table
freq_table = _create_frequency_table(text)
```

2. **Tokenization**: Breaking the text into sentences and words.

```python
# Tokenize
sentences = sent_tokenize(text)
```

3. **Sentence Scoring**: Calculating the score of each sentence based on word frequencies within it.

```python
# Score Sentences
sentence_scores = _score_sentences(sentences, freq_table)
```

4. **Selection of Best Sentences**: Selecting the sentences with the highest scores as the summary.

```
# Generate Summary
summary = _generate_summary(sentences, sentence_scores, 1.0 * threshold)
```

TextRank is an algorithm adapted from Google's PageRank to measure the importance of sentences within a text. TextRank works by building a sentence graph, where each sentence is treated as a node, and relationships between nodes are calculated based on sentence similarity.

The main steps in TextRank are:

1. **Sentence Graph Formation**: Sentences are connected based on content similarity.

```
#Tokenize Text
sentences = sent_tokenize(text)

#Similarity matrix
sentence_similarity_martix = _build_similarity_matrix(sentences, stop_words)

#Graph from similarity matrix
sentence_similarity_graph = nx.from_numpy_array(sentence_similarity_martix)
```

2. **PageRank Calculation**: Using the PageRank algorithm to calculate the importance of each sentence.

```
#PageRank
scores = nx.pagerank(sentence_similarity_graph)
```

3. **Selection of Important Sentences**: Sentences with the highest PageRank values are chosen as the summary.

```
#Ranked Sentence
ranked_sentence = sorted(((scores[i], s) for i, s in enumerate(sentences)), reverse=True)

#Top Sentence
for i in range(top_n):
    summarize_text.append(ranked_sentence[i][1])

#Summarize
return " ".join(summarize_text)
```

Testing and Evaluation

In NLTK, testing is done by calculating the frequency data within each sentence after removing stopwords and applying stemming, assigning scores and calculating average scores.

```
stopWords = set(stopwords.words("english"))
words = word_tokenize(text_string)
ps = PorterStemmer()

freqTable = dict()
for word in words:
    word = ps.stem(word)
    if word in stopWords:
        continue
    if word in freqTable:
        freqTable[word] += 1
    else:
        freqTable[word] = 1
return freqTable
```

```python
sentenceValue = dict()

for sentence in sentences:
    word_count_in_sentence = len(word_tokenize(sentence))
    word_count_in_sentence_except_stop_words = 0

    for wordValue in freqTable:
        if wordValue in sentence.lower():
            word_count_in_sentence_except_stop_words += 1
            if sentence[:10] in sentenceValue:
                sentenceValue[sentence[:10]] += freqTable[wordValue]
            else:
                sentenceValue[sentence[:10]] = freqTable[wordValue]

    if sentence[:10] in sentenceValue:
        sentenceValue[sentence[:10]] = sentenceValue[sentence[:10]] / word_count_in_sentence_except_stop_words

return sentenceValue
```

To measure the similarity of sentences using the text rank method, cosine similarity can be used. After the news text is written, the next step is to count the occurrences of words in the sentence and calculate the cosine similarity as in the following code :

```python
for w in sent_a:
    if w in stopwords:
        continue
    vector1[all_words.index(w)] += 1

for w in sent_b:
    if w in stopwords:
        continue
    vector2[all_words.index(w)] += 1

# cosine similarity
cosine_similarity = np.dot(vector1, vector2) / (np.linalg.norm(vector1) * np.linalg.norm(vector2))
return cosine_similarity
```

To evaluate the performance of both methods, we will examine a news summary generated using NLTK and TextRank. In this trial, we use several news articles varying in length and content structure. Below are the results and discussion of the implementation of both methods:

Original News Text:

This is the news text that will be summarized using the following two methods:

```python
text_str = '''
Presiden Jokowi dan Ibu Iriana Hadiri Pelantikan Presiden dan Wakil Presiden RI Masa Jabatan 2024-2029
Di sepanjang perjalanan menuju Gedung MPR/DPR, masyarakat juga tampak antusias menyambut iring-iringan Presiden dan Ibu Iriana,
melambaikan tangan sebagai bentuk penghormatan dan dukungan. Presiden Joko Widodo beserta Ibu Iriana Joko Widodo menghadiri pelantikan
Presiden dan Wakil Presiden Republik Indonesia masa jabatan 2024-2029 di Gedung MPR/DPR RI, Jakarta, pada Minggu, 20 Oktober 2024
Presiden Joko Widodo beserta Ibu Iriana Joko Widodo menghadiri pelantikan Presiden dan Wakil Presiden Republik Indonesia masa jabatan 2024-2029
di Gedung MPR/DPR RI, Jakarta, pada Minggu, 20 Oktober 2024. Presiden Jokowi dan Ibu Iriana bertolak dari Istana Merdeka Jakarta sekitar pukul 09.30 WIB menuju gedung parlemen.
Presiden Jokowi tampak mengenakan beskap hitam yang dipadu dengan kain nusantara, serta mengenakan peci hitam. Sementara itu,
Ibu Iriana tampak anggun dalam balutan pakaian nasional berwarna krem yang menambah kesan khidmat dalam momen bersejarah ini.
Sebelum meninggalkan Istana Merdeka, keduanya menyempatkan diri untuk menyapa para pegawai di lingkungan Istana Kepresidenan
Jakarta yang berbaris rapi di halaman istana. Di sepanjang perjalanan menuju Gedung MPR/DPR, masyarakat juga tampak antusias menyambut iring-iringan Presiden dan Ibu Iriana,
melambaikan tangan sebagai bentuk penghormatan dan dukungan.
Setibanya di Gedung MPR/DPR, Presiden Jokowi dan Ibu Iriana disambut hangat oleh Presiden terpilih Prabowo Subianto.
Turut menyambut pula Wakil Presiden Ma'ruf Amin dan Ibu Wury Ma'ruf Amin, Wakil Presiden terpilih Gibran Rakabuming Raka beserta Ibu Selvi Ananda,
serta Ketua MPR Ahmad Muzani beserta Ibu Himmatul Aliyah. Sebelum memasuki ruang sidang, Presiden Jokowi, Ibu Iriana, serta para pemimpin dan
pejabat negara yang hadir melangsungkan sesi foto bersama. Momen ini menjadi simbol solidaritas dan persatuan nasional dalam menyongsong era kepemimpinan baru.
Usai sesi foto, Presiden Jokowi, Ibu Iriana, dan seluruh tamu undangan memasuki ruang sidang untuk mengikuti prosesi pelantikan Presiden dan
Wakil Presiden Republik Indonesia masa jabatan 2024-2029. Sidang pun dimulai, menandai awal baru bagi pemerintahan Indonesia yang akan datang.
Turut hadir dalam acara pelantikan Presiden dan Wakil Presiden masa jabatan 2024-2029 tersebut antara lain sejumlah pemimpin dan utusan khusus negara sahabat,
Presiden ke-6 RI Susilo Bambang Yudhoyono, Wakil Presiden ke-6 RI Try Sutrisno, Wakil Presiden ke-10 dan ke-12 RI Jusuf Kalla,
serta Wakil Presiden ke-11 RI Boediono. Selain itu, tampak hadir pula para pimpinan lembaga negara dan para ketua umum partai politik nasional.
'''
```

Summary with NLTK:

Here is the summary generated using the NLTK method, with a summary length of 198 words from the original 376 words.

```
 Presiden Joko Widodo beserta Ibu Iriana Joko Widodo menghadiri pelantikan
Presiden dan Wakil Presiden Republik Indonesia masa jabatan 2024-2029 di Gedung MPR/DPR RI, Jakarta, pada Minggu, 20 Oktober 2024

Presiden Joko Widodo beserta Ibu Iriana Joko Widodo menghadiri pelantikan Presiden dan Wakil Presiden Republik Indonesia masa jab
atan 2024-2029
di Gedung MPR/DPR RI, Jakarta, pada Minggu, 20 Oktober 2024. Presiden Jokowi dan Ibu Iriana bertolak dari Istana Merdeka Jakarta
sekitar pukul 09.30 WIB menuju gedung parlemen. Presiden Jokowi tampak mengenakan beskap hitam yang dipadu dengan kain nusantara,
 serta mengenakan peci hitam. Sebelum meninggalkan Istana Merdeka, keduanya menyempatkan diri untuk menyapa para pegawai di lingk
ungan Istana Kepresidenan
Jakarta yang berbaris rapi di halaman istana. Di sepanjang perjalanan menuju Gedung MPR/DPR, masyarakat juga tampak antusias meny
ambut iring-iringan Presiden dan Ibu Iriana, melambaikan tangan sebagai bentuk penghormatan dan dukungan. Setibanya di Gedung MPR
/DPR, Presiden Jokowi dan Ibu Iriana disambut hangat oleh Presiden terpilih Prabowo Subianto. Sebelum memasuki ruang sidang, Pres
iden Jokowi, Ibu Iriana, serta para pemimpin dan
pejabat negara yang hadir melangsungkan sesi foto bersama. Usai sesi foto, Presiden Jokowi, Ibu Iriana, dan seluruh tamu undangan
 memasuki ruang sidang untuk mengikuti prosesi pelantikan Presiden dan
Wakil Presiden Republik Indonesia masa jabatan 2024-2029.
```

Summary with TextRank:

Here is the summary generated using the NLTK method, with a summary length of 65 words from the original 376 words.

```
Ringkasan:

Presiden Jokowi dan Ibu Iriana Hadiri Pelantikan Presiden dan Wakil Presiden RI Masa Jabatan 2024-2029
Di sepanjang perjalanan menuju Gedung MPR/DPR, masyarakat juga tampak antusias menyambut iring-iringan Presiden dan Ibu
Iriana, melambikan tangan sebagai bentuk penghormatan dan dukungan. Usai sesi foto, Presiden Jokowi, Ibu Iriana, dan sel
uruh tamu undangan memasuki ruang sidang untuk mengikuti prosesi pelantikan Presiden dan Wakil Presiden Republik Indones
ia masa jabatan 2024-2029.
```

Analysis of Results

1. **Cohesion and Coherence**
   TextRank: TextRank produces a more coherent summary by maintaining the flow between sentences. Since it uses a sentence graph structure, TextRank preserves the context of the news, making the summary appear more natural and sequential.
   NLTK: Although concise, the NLTK summary can sometimes lack coherence. NLTK selects sentences solely based on word frequency without considering inter-sentence relationships, which can make the summary feel somewhat disjointed.
2. **Information Relevance**
   TextRank: TextRank extracts more comprehensive information, covering main points like statements from the Minister of Energy and the President's periodic evaluation plans, showing TextRank's ability to capture the core of the news.
   NLTK: While NLTK captures some key points, it may miss crucial context, such as the President's policy evaluation and the main reason for the policy. This is due to NLTK's reliance on word frequency scores, which do not account for contextual connections between sentences.
3. **Speed and Performance**
   NLTK: Summarizing with NLTK is faster, as it only uses word frequency without building a sentence graph, making NLTK more efficient for processing very long documents.
   TextRank: TextRank takes longer because of the graph-building and PageRank calculation process. However, for relatively short news texts, the difference is negligible, making TextRank a viable option for this application.
4. **Summary Length**
   NLTK: NLTK's summaries are generally shorter, as it directly selects a few significant sentences without considering sentence connections.
   TextRank: TextRank summaries are slightly longer but still effective, as the method selects contextually relevant sentences.

CONCLUSION

TextRank demonstrates better performance in generating relevant and coherent news summaries compared to NLTK. The TextRank algorithm is capable of maintaining sentence relationships, which is crucial in news contexts to ensure information is conveyed accurately and fully. However, if efficiency is the top priority, as in very long texts, NLTK can be a quicker and more efficient choice, albeit at the expense of information cohesion.

Selecting the appropriate summarization method depends on the text's characteristics and the user's needs, while also considering advancements in summarization technology. NLTK and TextRank offer different approaches to extractive summarization, with NLTK being more suitable for preprocessing and TextRank excelling at selecting contextually important sentences. TextRank performs well on news texts with compact and dense sentence structures due to its sentence-ranking capability. Meanwhile, NLTK serves as a good supplement in the initial stages of text processing.

The results indicate that TextRank's summaries are more informative and contextual for news, whereas NLTK is more focused on basic text processing efficiency. Combining the two can produce a more optimal summary. This study demonstrates that Python-based news text summarization is easy to implement, and in the future, hybrid methods or deep learning models like BERT and T5 could significantly enhance text summarization quality.

REFERENCES

Alami Merrouni, Z., Frikh, B., & Ouhbi, B. (2023). EXABSUM: a new text summarization approach for generating extractive and abstractive summaries. *Journal of Big Data*, *10*(1). https://doi.org/10.1186/s40537-023-00836-y

Albab, M. U., Karuniawati P, Y., & Fawaiq, M. N. (2023). Optimization of the Stemming Technique on Text preprocessing President 3 Periods Topic. *Jurnal TRANSFORMATIKA*, *20*(2), 1–10. https://journals.usm.ac.id/index.php/transformatika/■page1

Barrios, F., López, F., Argerich, L., & Wachenchauzer, R. (2016). Variations of the Similarity Function of TextRank for Automated Summarization. *Argentine Symposium on Artificial Intelligence (ASAI) 2015*. http://arxiv.org/abs/1602.03606

Hernawan, Y. F., Adikara, P. P., & Wihandika, R. C. (2022). Peringkasan Artikel Berbahasa Indonesia Menggunakan Textrank Dengan Pembobotan BM25. *Jurnal Teknologi Informasi Dan Ilmu Komputer (JTIIK)*, *9*(1), 61–68. https://doi.org/10.25126/jtiik.202293765

Jimly Hanif, A., Nur Farid, M., & Hasanah, B. (2023). Penerapan Natural Language Processing untuk Klasifikasi Bidang Minat berdasarkan Judul Tugas Akhir. *Jurnal Sistim Informasi Dan Teknologi*, *5*(1), 41–49. https://doi.org/10.37034/jsisfotek.v5i1.196

Kumar, T. R., Akarapu, A., Allam, S. R., Gattepally, N. C., & Nellutla, P. yashaswi. (2024). Chatbot Application Using Nltk and Keras. *International Research Journal of Engineering and Technology (IRJET)*, *11*(04), 2540–2548. www.irjet.net

Louis, D. A., Rostianingsih, S., & Santoso, L. W. (2022). Implementasi Text Summarization pada Review Aplikasi Super di Google Play Store Menggunakan Metode Maximum Marginal Relevance. *Jurnal Infra*, *10*(2), 186–190.

Mahawardana, P. P. O., Sasmita, G. A., & Pratama, I. P. A. E. (2022). Analisis Sentimen Berdasarkan Opini dari Media Sosial Twitter terhadap "Figure Pemimpin" Menggunakan Python. *JITTER- Jurnal Ilmiah Teknologi Dan Komputer*, *3*(1). https://ojs.mahadewa.ac.id/index.php/jmti/article/view/2111

Munthe, M. (2022). Penerapan Metode Textrank dalam Rancangan Aplikasi Silogisme Artikel Bahasa Batak. *Journal of Computing and Informatics Research*, *1*(2), 37.

Pertiwi, L. (2022). Penerapan Algoritma Text Mining, Steaming Dan Texrank Dalam Peringkasan Bahasa Inggris. *BIMASATI(Bulletin of Multi-Disciplinary Science and Applied Technology)*, *1*(3), 100–104.

Rifano, E. J., Fauzan, A. C., Makhi, A., Nadya, E., Nasikin, Z., & Putra, F. N. (2020). Text Summarization Menggunakan Library Natural Language Toolkit (NLTK) Berbasis Pemrograman Python. *ILKOMNIKA:*

Journal of Computer Science and Applied Informatics, *2*(1), 8–17. https://doi.org/10.28926/ilkomnika.v2i1.32

Sadam, E. F., Annafi, A. R., Prasetio, I. W., Nashiro, M. S., Septyawan, R., Sutrisno, M., Risanty, R. D., Ambo, S. N., Adharani, Y., & Mujiastuti, R. (2024). Basic Introduction to NLP : Text and Language Analysis with Python. *Jurnal Pendidikan Tambusai*, *8*(2), 34975–34982.

Setiawan, A. Y. S. A., & Alexander, E. (2023). Pengujian Algoritma TextRank Dalam Merangkum Teks. *Seminar Nasional Amikom Surakarta (Semnasa) 2023*, *November*, 25–35.

Setiawan, A. Y. S. A., & Alexander, E. (2024). PENERAPAN ALGORITMA TEXTRANK DALAM MERANGKUM TEKS WORD DAN P. *JITET (Jurnal Informatika Dan Teknik Elektro Terapan)*, *12*(1).

Umunyana, C., Tuyizere, G., & Mbarushimana, A. (2024). Students' Sentiment Analysis Using Natural Language Toolkit in Machine Learning for Module Evaluation. *European Journal of Engineering and Technology Research*, *9*(1), 72–75. https://doi.org/10.24018/ejeng.2024.9.1.3006

Wang, M., & Hu, F. (2021). The Application of NLTK Library for Python Natural Language Processing in Corpus Research. *Theory and Practice in Language Studies*, *11*(9), 1041–1049. https://doi.org/10.17507/tpls.1109.09