



PROCEEDING

2022 International Symposium on Information Technology and Digital Innovation (ISITDI)

Virtual Symposium July 27-28, 2022

Theme

Technology Innovation During Pandemic

2022 International Symposium on Information Technology and Digital Innovation (ISITDI) | ISBN: 978-1-6654-6116-0 | DOI: 10.1109/ISITDI57943.2022.9944449

ISBN: 978-1-6654-6116-0
Part Number: CFP22CE7-ART

© Copyright 2022 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to use any copyrighted component of this work in other work must be obtained from the IEEE

ORGANIZED BY:



TECHNICAL CO-SPONSORED BY:



**2022
International Symposium on Information Technology and
Digital Innovation
(ISITDI)**

July 27-28, 2022

Copyright and Reprint Permission:

Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through:

Copyright Clearance Center
222 Rosewood Drive
Danvers
MA 01923.

For reprint or republication permission, email to:

IEEE Copyrights Manager at pubs-permissions@ieee.org

All rights reserved.

Copyright ©2022 by IEEE.

IEEE Catalog Number CFP22CE7-ART

ISBN 978-1-6654-6116-0

PROCEEDING

2022
International Symposium on Information
Technology and Digital Innovation
(ISITDI)

July 27-28, 2022
(Virtual Symposium)

PREFACE

We are delighted to welcome you virtually to the 2022 International Symposium on Information Technology and Digital Innovation (ISITDI). The theme for this year is ***Technology Innovation During Pandemic***. The symposium is held by the Faculty of Information and Technology, the Universitas Andalas, Indonesia, and the Universitas Dharma Andalas, Indonesia.

Furthermore, this symposium is divided into eight topics: *Services Computing and Services Computing Systems, Smart Services Systems and Engineering, Digital Security, Forensic and Cyber, Digital Transformation and Innovation, IT Governance and Management, Internet of Things, Computing and Engineering, and Miscellaneous*. We hope that all the participants can capitalize on this event and gain benefits from it. A total of 37 papers were accepted, consisting of authors distributed from Indonesia, Japan, Malaysia, and Bangladesh.

In closing, I would like to thank the Rector of Universitas Andalas and Universitas Dharma Andalas and also all committee members of this symposium for their hard work and patience.

Dr.Eng. Budi Rahmadya (Universitas Andalas, Indonesia)

Conference Chair of ISITDI 2022

Committee

Conference Chair	Dr. Eng. Budi Rahmadya
Vice Chair of the Conference	Afriyanti Dwi Kartika, S.Pd, M.T
Technical Program Committee	Dr. Fakir Sharif Hossain
	Dr. Rian Ferdian, MT
	Prof. Dr. H. Jufriadif Na'am
	Dr. Eng. Muhammad Ilhamdi Rusydi
	Dr. Eng. Agussalim
	Dr. Eng. Budi Irmawati
	Dr. Eng. Muhammad Ashar
	Ade Irawan Ph.D
Publishing Chair	Hafizah Hanim, M.Kom
	Nefy Puteri Novani, MT
Secretary	Rifki Suwandi, MT
Organizing Committee	Rizka Hadelina, MT
	Desta Yolanda, MT
	Jefril Rahmadoni, M.Kom
	Ricky Akbar, M.Kom
	Arrya Anandika, MT
	Iswan Rina, M.Si
	Dwi Sulistiowati, M.Si
	Nurweni Puteri, M.Si
	Maya Sari Syahrul, M.Si
	Renita Astri, S.Kom, M.Sc
	Zulfahmi, S.Kom, M.Kom
	Faradika, S.Kom, M.Kom
	Putri Anggraini, S.Kom, M.Kom
	Sularno, S.Kom, M.Kom
	Muhammad Razi A, ST, M.M.SI
	Dio Prima Mulya, S.Kom, M.Kom

Table of Content

Session Number	Title Authors	Page
R1-01	Gender and Intent Classification from Finger Swiping Behaviours on Gesture Keyboards Using LSTM <i>Ryan Adipradana, Bernard Wijaya Winston, Rusli Henry Lucky, and Derwin Suhartono</i>	1 - 5
R1-02	An RFID-Based Battery-Less Vibration Monitoring System for Electrical Appliances <i>Zequan Song, Ran Sun, Budi Rahmadya, and Shigeki Takeda</i>	6 - 10
R1-03	Image Enhancement on Deep Learning Algorithm for COVID-19 Lung X-Ray Classification <i>Said Thaufik Rizaldi, Mustakim, Inggih Permana, and M Afdal</i>	11 - 15
R1-04	Comparison of Naïve Bayes, C4.5 and K- Nearest Neighbor for Covid-19 Data Classification <i>Umairah Rizkya Gurning, Mustakim, and Said Thaufik Rizaldi</i>	16 - 21
R1-05	Fatigue Warning System During Physical Exercise Based on Heart Rate and Oxygen Saturation Using Non-invasive Wearable Sensor <i>Dody Ichwana and Shelvi Ekarian</i>	22 - 26
R1-06	The Implementation Of Hand Gestures On The Fountain Burst Pattern <i>Rahmi Eka Putri and Nur Afni Nazara</i>	27 - 31
R1-07	LPWAN Communication in IoT Network for Electrical Energy Monitoring <i>Hanalde Andre, Rizki Wahyu Pratama, Febi Sabila, Yoppi Lisyadi Oktapianus, Teddy Yuliswar, Darmawan, Dwi Welly Sukma Nirad, Andrew Kurniawan Vadreass, and J. Rahmadoni</i>	32 - 35
R2-01	Designing Service Architecture Platform using SCSE Framework for National Archive System Based on OAIS and Microservice Modeling <i>Ninon Nurul Faiza, Suhardi, and Taufik Iqbal Ramdhani</i>	36 - 41
R2-02	The Implementation of Business Intelligence on Visualisation of Transaction Data Analysis using Dashboard System Case Study: XYZ Convenience Store <i>Ricky Akbar, Wahyudi, J. Rahmadoni, and Hanifah Lubis</i>	42 - 49
R2-03	Gamification Methods of Game-Based Learning Applications in Medical Competence: A Systematic Literature Review	50 - 54

Session Number	Title Authors	Page
	<i>Lazuardi Fatahilah Hamdi, Bimo Sunarfri Hantono, and Adhistya Erna Permanasari</i>	
R2-04	Notification System and GPS Position Tracking as a Security Feature for Child Pick Up at Daycare <i>Ratna Aisuwarya, Tati Erlina, Syafajar Ahmad Sabyl</i>	55 - 58
R2-05	Impact Big Data Analytics for Healthcare Sector Using Classification Algorithm <i>Anita Sindar Ros Maryana Sinaga</i>	59 - 64
R2-06	Coffee Shop Recommendation System Using An Item-Based Collaborative Filtering Approach <i>Renita Astri, Ahmad Kamal, and Suaini Sura</i>	65 - 67
R3-01	Asset Growth Classification Based on Financial Performance of LQ45 Indexed Companies by Using Logistic Regression <i>Nurweni Putri, Dodi Devianto, and Tia Hafazah</i>	68 - 73
R3-02	Social Media Influence on Effectiveness of Customer Knowledge Management: Case study of PT XYZ <i>Rommy Bastian Hutauruk</i>	74 - 79
R3-03	Essay Test Based E-Testing Using Cosine Similarity Vector Space Model <i>Wahyudi Wahyudi, Ricky Akbar, Teguh Suharsono, and Ahmad Indrapriyatna</i>	80 - 85
R3-04	Detection, Prevention and Emergency Solution of Road Accidents in Bangladesh using IoT <i>Md. Sumon Fakir, Najmus Sakib, Md. Solaiman Mia, and Najmus Sakib Sizan</i>	86 - 91
R3-05	Usability Evaluation Methods of Mobile Applications: A Systematic Literature Review <i>Adi Nugroho, Paulus Insap Santosa, and Rudy Hartanto</i>	92 - 95
R3-06	Design of Socio-Technopreneur Platform Model as Innovative Solution to Environmental and Economic Problems <i>Dwi Welly Sukma Nirad, Afriyanti Dwi Kartika, Andrew Kurniawan Vadreas, Hanalde Andre, Winda Amelia, and Salsabila Rahmah</i>	96 - 100
R4-01	Development of Component Recognition Applications and Labor Tools Based On Android and Tiny Yolo Networks (Case Study: Signal and System Laboratory) <i>Dodon Yendri, Desta Yolanda, and Lathifah Arief</i>	101 - 107
R4-02	Automated Text Summarization and Topic Detection on News Aggregation System Using BART and SVM	108 - 113

Session Number	Title Authors	Page
	<i>Albert Wihardi, Farrel Octavianus, Muhamad Keenan Ario, and Derwin Suhartono</i>	
R4-03	Sexual Violence Classification as Hate Speech using Indonesian Tweet <i>Muhammad Notareza Ramadhan, Indra Budi, Aris Budi Santoso, and Ryan Randy Suryono</i>	114 - 120
R4-04	Application of Residual Network Architecture on Covid-19 Chest x-ray Classification <i>Susanti Susanti, Mustakim Mustakim Rice Novita, and Inggih Permana</i>	121 - 125
R4-05	Two-Way Multivariate Analysis of Variance in Comparative Analysis of Study Exam Scores Based on Learning Methods and Student Gender <i>Dwi Sulistiowati, Iswan Rina, and Nadia Febrian</i>	126 - 130
R4-06	Graph Coloring Applications in Scheduling Courses using Welch-Powell Algorithm - A case study <i>Iswan Rina and Dwi Sulistiowati</i>	131 - 135
R4-07	COVID-19 SEIRV Model Simulation In West Sumatra Province Using Runge Kutta 4th Order <i>Maya Sari Syahrul</i>	136 - 139
R4-08	Hidden Markov Model as the Predictive Time Series Movement of Natural Gas Price <i>Dodi Devianto</i>	140 -145
R4-09	Identification of Motorcycle Traffic Violations with Deep Learning Method <i>Rian Ferdian and Tiara Permata Sari</i>	146 - 149
R5-01	Strategy to Improve Data Quality Management: A Case Study of Master Data at Government Organization in Indonesia <i>Rizqa Nulhusna, Nur Fajar Taufiq, and Yova Ruldeviyani</i>	150 - 155
R5-02	BERT and ULMFiT Deep Ensemble Model for Personality Prediction from Social Media Text <i>Noptovius Halimawan, Derwin Suhartono, Aryo Gema, and Rezki Yunanda</i>	156 - 161
R5-03	Network Automation for CE Router with Route Leaking in MPLS-VPN Network <i>Mochamad Yazid Gupron, Umaisaroh, and Agung Wicaksono</i>	162 - 167
R5-04	A Safe Approach to Sensitive Dropout Data Collection Systems by Utilizing Homomorphic Encryption <i>Rifki Suwandi and Aciek Ida Wuryandari</i>	168 - 171

Session Number	Title Authors	Page
R5-05	Design and Analysis of Bandung-Jakarta High-Speed Rail Fiber Optic Backbone Network <i>Naufal Inas Fikri, Primayoga Budyprawira, Listi Farida, Stevan Pasaribu Muhammad Athalah, Reza Febrian, and Catur Apriono</i>	172 - 176
R5-06	Security and Privacy Issue in Internet of Things, Smart Building System: A Review <i>Inggit Putri Naria, Selo Sulistyono, and Widy Widyawan</i>	177 - 180
R5-07	Smart City Service System Design Based On Microservice Architecture: Case Study in Magelang City <i>Muhammad Tri Adhi Utomo, Rudy Hartanto, and Selo Sulistyono</i>	181 - 187
R5-08	HRIS Implementation Process: Case Study on Bank XYZ <i>Anggoro Gagah Nugroho, Haryani Diah Sitawati, Hidayat Akbar, Retiana Fatma Pertivi, and Panca O. Hadi Putra</i>	188 - 194
R5-09	Development of the Minangkabau Local Language Translation Machine Based on Stemming <i>Rini Sovia</i>	195 - 198

Development of the Minangkabau Local Language Translation Machine Based on Stemming

Rini Sovia
University of Putra Indonesia YPTK
Padang, Indonesia
rini_sovia@upiypk.ac.id

Sarjon Defit
University of Putra Indonesia YPTK
Padang, Indonesia
sarjond@yahoo.co.uk

Yuhandri
University of Putra Indonesia YPTK
Padang, Indonesia
yuyu@upiypk.ac.id

Abstract— Indonesia is an archipelagic country that has hundreds of ethnic groups and regional languages. One of the well-known regional languages is the Minangkabau language (BM) which is dominantly used in several areas in Sumatra which are in the Austronesian family. The habit of the Minangkabau people in their daily life is to always use the Minangkabau language (BM) in communicating. Usually, the Minang tribe always communicates every day using the Minangkabau language, so that it is unique for the people around them, thus creating curiosity to know BM. So this research was conducted to translate BM into Indonesian. The purpose of this research is to translate BM into Indonesian. By using the translation engine of the Minangkabau Language Stemming Algorithm (SBMK). The data processed were 600 basic words in printed dictionaries and sentences in 12 BM documents. The level of accuracy of the translation results from this study is 98.33% for basic words and 94.68% for sentences in the document. The resulting algorithm is very precise to translate and process the basic word spelling checker in BM words and documents into Indonesian.

Keywords—*language, Minangkabau, translation engine, spelling checker, basic words.*

I. INTRODUCTION

The Minangkabau language is a regional language used by the Minangkabau people from the Minangkabau Highlands in West Sumatra, South Sumatra, and the west coast of the Mukomuko region [1]. The Minangkabau language (BM) is very popular with its various dialects, such as Agam, Batu Sangkar, Pesisir, Solok, and Pariaman [2]. BM has several unique words in prefixes, insertions, suffixes, combinations, and disconnected affixes. In prefixes consisting of ba-1, ba-2, maN, paN-, pa-, ta, no, di, sa, ka, raw, and basi, insert -il, -al, -ar, -am, and ij, endings -an, -kan, I, and -lah, compound ba-Kan, ba-1, no-Kan, pa-Kan, ba-lah, standar-lah, stale-lah, man-pa-Kan, no-pa-Kan, no-sa-Kan, sa-paN, di-pa-sa-Kan, and interrupted affixes Ka.an, Ka.no, paN.an. The uniqueness that exists in the Minangkabau language is in the insertion affix, where of the five insertion words in Minangkabau language, only il, ar, am are widely used, which are not too productive[3]. Morphologically rich language, and morphological and ambiguous analysis plays an important role in most Natural Language Processing (NLP) tasks[4]. NLP is a branch of artificial intelligence that focuses on natural language processing. The language that would be understood by the computer requires a process first so that the user's wishes can be clearly understood by the computer[5]. Stemming is a sub-field of NLP, which is a phase process in pre-processing finding the root or root word in a particular word [6][7][8]. Stemming is widely used in application development, especially in terms of Information Retrieval (IR), and text mining, to improve system performance [9][10][11]. The stemming function here is to cut or separate

the basic words with affixes, both prefixes, insertions, suffixes, or combinations [12], [13].

The stemming algorithm had widely used for many cases, such as determining similarities in submitting thesis titles using Nazief & Adriani stemming [14]. Then compare two Indonesian stemmers Porter and Arifin Setiono, to find out which stemmer is more effective in determining the root word [15]. Arifin and Setiono also proposed a new algorithm similar to Nazief, but adding affixes to words to be omitted, resulting in a more effective root word [16]. Stemming on tweet documents to analyze the public opinion of Indonesian tweets about presidential candidates of the Republic of Indonesia in 2014 using Naive Bayes classification, Maximum Entropy classification, and Support Vector Machines[17]. ECS stemming reduces the number of terms generated at the preprocessing stage by using the Clustering method[18]. Affix grouping based on Indonesian morphology stemming algorithm Enhanced Confix Stripping (ECS), New Enhanced Confix Stripping (NECS) stemming algorithm, and UG18 stemming algorithm[19].

The stemming methods that exist in each language are different from each other, where Indonesian stemming has a different morphology from the Minangkabau language stemming. Stemming for the Minangkabau language is more complicated because several affixes will be removed to get the root word. Stemming regional languages using the Rule-Based Approach which produces an accuracy rate of 96.94% with a total of 120 incorrect words corrected to 20 incorrect words[20], modification of the Enhanced Confix Stripping stemmer method, using data in the form of text/poetry in the Madurese language [21].

II. MATERIAL AND METHOD

A. The Spell Check

Spell Check is the process of checking for spelling errors of words in the text and providing solutions for errors automatically. Errors that arise can be caused by the use of the wrong words, and typing and coding errors. Spelling errors are divided into two, namely non-word errors and real-word errors. Non-word errors occur because the typed word is not in the dictionary, the word is in the dictionary but is wrong in the context [22], [23]. The challenges in making a spelling checker are in the process of finding the wrong word and providing suggestions in the form of the right word to replace the mistake word, as well as the process of recognizing grammar in sentences, whether ambiguity and words that do not exist in the dictionary are also known as Out of Vocabulary (OOV) While errors in non-words, the process of checking excessive letters and spelling words [24].

B. Minangkabau language (BM)

Minangkabau language (BM) has three types of word meanings (phonemes) [25]. The three phonemes are 5 vowels, namely a, i, u, e, and o; 20 consonants, and 6 diphthongs, namely id, ud, aw, ay, uy, ed [26]. The smallest words (morphemes) in BM consist of 1 to 4 syllables that have meaning [27]. Morphological morpheme processes are grouped into seven groups of affixes which are presented in TABLE I.

TABLE I BM AFFIX GROUP

No	Group	Affix
1.	Prefix	<i>ba-I, ba-2, maN, paN-, pa-, ta, no, sa, baku, baka, basi, ka, bapa, tapa, maN pa, sa pa</i>
2.	Insert	<i>-il, -al, -ar, -am, ij</i>
3.	Suffix	<i>-an, -kan, i, dan lah</i>
4.	Disconnected Affix	<i>ka..an, ka..no, paN..an</i>
5.	Combination of prefix and suffix	Combination of prefix and suffix (<i>ba Kan, ba a- i, no- Kan, pa- Kan, ba- lah, baku- lah, ba si- lah</i>), Combind Suffix and Prefix (<i>MaN- pa- Kan, no- pa- Kan, No- sa- Kan, sa- paN, di- pa- s a- Kan</i>)
6.	Combination of prefix and combination of suffix	<i>maN..pa..Kamlah, maN..sa..Kamlah dipa..Kamlah, disa..Kamlah, baku..lah, basi.. lah, sapaN..lah</i>
7.	Other disconnected affixes	<i>ba2..ka..an, ba2..paN..an, sa..paN..an</i>

Based on the group of affixes in TABLE I, word formation can be done using (1).

$$kd = [aw] + [akh] + [dk] + [sis] + gab \quad (1)$$

Where kd is a basic word in documents and sentences, aw is a prefix, akh is a suffix, ds is a basic word, sis is an insertion in a sentence, and gab is a combination of affixes. The part of the word that is combined with the root word will form an affix. The Minangkabau Language Stemming Algorithm (SBMK) process begins with finding the word to be stemmed in the dictionary. If a word is found, it becomes the root word and the process stops. If the word is not found, then the deletion process is carried out starting from the deletion of the prefix, the deletion of the suffix, the removal of the insert, the deletion of the interrupted affix, and the deletion of the combined. All processes refer to checking in the Minangkabau language dictionary. If the word you are looking for is not found in the dictionary, then the word you are looking for becomes the root word.

Documents containing variations in various forms of letters and punctuation, need to be uniformed through a preprocessing process with the aim that the data used is free of noise. The preprocessing stage includes case folding, tokenizing, stopword removing, and stemming processes[28]. Case Folding is the process of changing the entire text in a document into lowercase letters, such as 'a', 'b', etc., tokenizing is the process of separating a document into parts, and removing some characters, such as punctuation marks. Stopword Removing is the process of removing words that have no meaning, such as 'and', 'or', 'by'[29]. Stemming is the process of separating root words from prefixes (prefixes), insertions (infixes), suffixes (suffixes), and combinations (confixes)[30].

The stages of the process of checking the spelling and translation of the Minangkabau language are presented in Fig. 1.

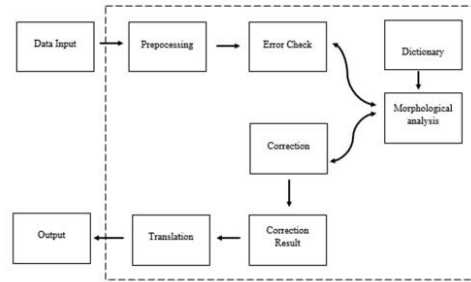


Fig. 1. Stages of the Translation Process

Fig. 2, describes the process carried out in checking the spelling of the Minangkabau language, starting from the preprocessing stage, before proceeding to the next stage, a language dictionary is needed to check words according to the morphological analysis of the language used. The preprocessing stage consists of processes, such as case folding, which removes all periods and punctuation marks in a document, then proceeds with the tokenizing process, which is the process of separating each syllable, then the stopword removing process, which removes words. words that have no meaning, such as the word and, or, by, etc. Then there is the stemming process, which is to remove existing affixes such as prefixes, insertions, and suffixes. Then proceed to the error detection and error correction process. After checking and correcting errors which refers to the analysis of the morphology of the language, then it produces results in the form of words in the document. Next, carry out the language translation process, according to the EYD rules in Indonesian. The algorithm of the translation process is presented in the following pseudocode in Fig. 2.

```

Translasi Algorithm
Input      : KD, Kata, Kal
Output     : KD, Kata, Kal
Initialization preg_match
If (cekKamus($1_KD)){
  $data['kata1']=$1_KD;
  $data['kata2']=$2_Kata;
  $data['kata3']=$3_Kal;
}
Else
  If (preg-match 'KD'){
    If (preg-match('KD')){
      Return Tampil KD;
      Return Arti KD;
    }
  }
  Else
    If (preg-match ('kata')){
      If (preg-match ('kata')){
        Return Tampil kata;
        Return Arti kata;
      }
    }
  }
  Else
    If (preg-match('Kal')){
      Return hapus Kal;
      Return Arti Kal;
    }
  }
End If
End If
  
```

Fig. 2. Translation Pseudocode

The translation algorithm in Fig. 1 is based on the grouping of basic words, words, and sentences in the morphology of the Minangkabau language. The translation process is carried out starting from the root word. This algorithm processes basic words, words, and sentences in the document which will produce basic words, words, and sentences in the Minangkabau language. Base words, words, and sentences will be validated with the database. Basic words, words, and sentences found in the database will be processed to produce basic words, words, and sentences that have been translated into Indonesian. Like the word, "barangkek" will be "depart".

III. RESULT AND DISCUSSION

The translation algorithm was tested on 600 basic words. The choice of words tested was based on the groups of vowels and consonants in the database. The translation algorithm was also tested on 12 Minangkabau language folklore documents. Each test result is validated by an expert and the formula to determine the level of accuracy in the word is presented in (2). Accuracy values for words that were successfully translated in the document using (3).

$$\text{Word Translation Acc} = \frac{\sum SW}{\sum IW} \times 100\% \quad (2)$$

$$\text{Doc Translation Acc} = \frac{\sum SD}{\sum ID} \times 100\% \quad (3)$$

Where $\sum SW$ is the number of successful word translations, and $\sum IW$ is the number of words tested. $\sum SD$ is the total translation of documents, and $\sum ID$ is the number of test documents. The algorithm application is implemented using the PHP Programming Language with test data in the form of a dictionary stored in a MySQL database. One of the test results using the application is presented in Fig. 4.

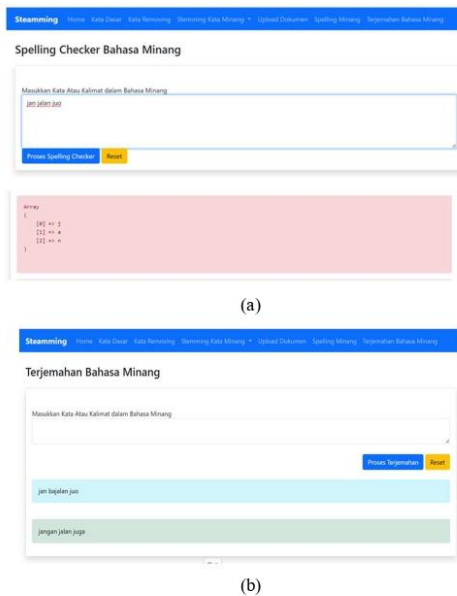


Fig. 4. Testing interface (a) Spelling Checker, (b) Translator results

TABLE II. TEST RESULTS ON THE WORD

No.	Group	Word Count	Word Translate	Accuracy (%)
1.	prefix	387	385	99.00
2.	insert	11	10	91.00
3.	suffix	96	93	97.00
4.	affix	59	57	97.00
5.	Combination of prefix and suffix	18	17	94.00
6.	Combind prefix and combined suffix	24	23	96.00
7.	Another disconnected affix	5	5	100.00
Total		600	590	
Average				98.33

TABLE III. TEST RESULTS ON WORDS IN THE DOCUMENT

No	Title	Word Count	Word Translate	Accuracy (%)
1.	Asal usul Maninjau.txt	199	190	95.00
2.	Mande.txt	72	65	90.00
3.	Cerita Minang.txt	112	100	89.00
4.	Mengutaroran Cinto.txt	1,403	1,320	94.00
5.	Barubek.txt	394	372	94.00
6.	Di rumah Puti Galang.txt	453	435	96.00
7.	Talarak dek harato.txt	1,722	1,700	99.00
8.	Mandapek Malu.txt	518	480	93.00
9.	Di tingga Marantau.txt	193	180	93.00
10.	pituah bapak jo mande.txt	477	464	97.00
11.	Malin Kundang.txt	399	350	88.00
12.	Marantau.txt	507	450	89.00
Total		6,449	6,106	
Average				94.68

Based on the test results in TABLE II and III, the average accuracy level of translators from the SBMK algorithm is obtained, namely:

$$\text{Accuracy} = \frac{\text{Word Trans Accu} + \text{doc Trans Accu}}{2} \quad (4)$$

$$= \frac{98.33\% + 94.68\%}{2} = 96.50\%$$

With an accuracy result of 96.50%, it makes the translation algorithm reliable, and has advantages in translating words, sentences in documents. Another advantage of the translation algorithm is that it can work very well and can also identify words and spelling checkers in sentences.

IV. CONCLUSION

The translation algorithm is a standard stemming algorithm for the Minangkabau language which can be implemented for translating words and sentences in documents. The translation algorithm can determine the spelling checker for words and sentences in the document. The

system produces an accuracy rate of 96.50% from 600 words and 12 documents containing.

REFERENCES

- [1] M. Nur, "King Pagaruyung in Minangkabau in Historical Perspective," *J. Anal. Sej.*, vol. 9, no. 2, pp. 9–29, 2020.
- [2] F. Rahman and S. Kumiati, "Tbhas : Ludling the Minangkabau dialect of the South Coastal language," *KEMBARA (Scientific Journal of Language, Literature, and Its Teaching)*, vol. 7, no. 2, pp. 51–64, 2021.
- [3] B. K. H. Nio, "Minangkabau Language Morphology and Syntax," *History*, vol. I, no. 1636, pp. 1631–1634, 1979.
- [4] A. Nzeyimana, "Morphological disambiguation from stemming data," pp. 4649–4660, 2021, doi: 10.18653/v1/2020.coling-main.409.
- [5] U. Hasanah, T. Astuti, R. Wahyudi, Z. Rifai, and R. A. Pambudi, "An experimental study of text preprocessing techniques for automatic short answer grading in Indonesian," *Proc. - 2018 3rd Int. Conf. Inf. Technol. Inf. Syst. Electr. Eng. ICTISEE 2018*, pp. 230–234, 2018, doi: 10.1109/ICTISEE.2018.8720957.
- [6] A. S. Rizki, A. Tjahyanto, and R. Trialih, "Comparison of stemming algorithms on Indonesian text processing," *Telkomnika (Telecommunication Comput. Electron. Control)*, vol. 17, no. 1, pp. 95–102, 2019, doi: 10.12928/TELKOMNIKA.v17i1.10183.
- [7] F. S. Utomo, N. Suryana, and M. S. Azmi, "Stemming impact analysis on Indonesian Quran translation and their exegesis classification for ontology instances," *IJUM Eng. J.*, vol. 21, no. 1, pp. 33–50, 2020, doi: 10.31436/iujumej.v21i1.1170.
- [8] C. Moral, A. de Antonio, R. Imbert, and J. Ramirez, "A survey of stemming algorithms in information retrieval," *Inf. Res.*, vol. 19, no. 1, 2014.
- [9] Y. Jaafar, D. Namly, K. Bouzoubaa, and A. Yousfi, "Enhancing Arabic stemming process using resources and benchmarking tools," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 29, no. 2, pp. 164–170, 2017, doi: 10.1016/j.jksuci.2016.11.010.
- [10] I. Shrestha and S. S. Dhakal, "A new stemmer for Nepali language," *Proc. - 2016 Int. Conf. Adv. Comput. Commun. Autom. (Fall)*, ICACCA 2016, pp. 0–4, 2016, doi: 10.1109/ICACCAF.2016.7749008.
- [11] H. A. R. Nur Hidayatullah, Aji Prasetya Wibawa, "Application of ECS Stemmer for Nazief & Adriani Modifications in Javanese," vol. 1, no. 10, 2019.
- [12] R. Setiawan, A. Kumiawan, W. Budiharto, I. H. Kartowisastro, and H. Prabowo, "Flexible affix classification for stemming Indonesian Language," *2016 13th Int. Conf. Electr. Eng. Comput. Telecommun. Inf. Technol. ECTI-CON 2016*, 2016, doi: 10.1109/ECTICon.2016.7561257.
- [13] I. P. M. Wirayasa, I. M. A. Wirawan, and I. M. A. Pradnyana, "Bastal Algorithm: Nazief & Adriani Algorithm Adaptation for Balinese Text Stemming," *J. Nas. Pendidik. Tek. Inform.*, vol. 8, no. 1, p. 60, 2019, doi: 10.23887/janapati.v8i1.13500.
- [14] H. Pramudita, "Application of Nazief & Adriani Stemming Algorithm and Similarity in Acceptance of Thesis Title," *Data Manaj. dan Teknol. Inf.*, vol. 15, no. 4, p. 15, 2014.
- [15] D. Novitasari, "Comparison of Porter's Stemming Algorithm with Arifin Setiono to Determine the Accuracy Level of Basic Words," *STRING (Unit of Writing Res. and Inov. Teknol.)*, vol. 1, no. 2, p. 120, 2017, doi: 10.30998/string.v1i2.1031.
- [16] A. S. Rizki, "Comparison of Indonesian Stemmers and Their Impact on Indonesian Text Mining, Case Study of PLN Customer Complaints Grouping," p. 205, 2017, [Online]. Available: <http://repository.its.ac.id/43254/>
- [17] B. H. Iswanto and V. Poerwoto, "Sentiment analysis on Bahasa Indonesia tweets using Unigram models and machine learning techniques," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 434, no. 1, pp. 0–6, 2018, doi: 10.1088/1757-899X/434/1/012255.
- [18] T. Winarti, J. Kerami, and S. Arief, "Determining Term on Text Document Clustering using Algorithm of Enhanced Confix Stripping Stemming," *Int. J. Comput. Appl.*, vol. 157, no. 9, pp. 8–13, 2017, doi: 10.5120/ijca2017912761.
- [19] I. Mulyana, A. Suhendra, Emastuti, and W. Bheta Agus, "Development of Indonesian stemming algorithms through modification of grouping, sequencing and removing of affixes based on morphophonemic," *Int. J. Recent Technol. Eng.*, vol. 8, no. 2 Special Issue 7, pp. 179–184, 2019, doi: 10.35940/ijrte.B1044.07825719.
- [20] N. W. Wardani and P. G. S. C. Nugraha, "Balinese Text Stemming with Enhanced Confix Stripping Algorithm," *Int. J. Nat. Sci. Eng.*, vol. 4, no. 3, p. 103, 2020, doi: 10.23887/ijnse.v4i3.30309.
- [21] R. Maulidi, "Modification of Enhanced Confix Stripping Method," *Pros. Semin. Nas. FDI 2016*, no. December, pp. 12–15, 2016.
- [22] S. H. Wibowo, B. Soerowirdjo, Ernastuti, and A. Tarigan, "Spelling checker of words in rejang language using the n-gram and euclidean distance methods," *J. Comput. Theor. Nanosci.*, vol. 16, no. 12, pp. 5384–5395, 2019, doi: 10.1166/jctn.2019.8607.
- [23] U. Liyanapathirana, K. Gunasinghe, and G. Dias, "Sinspell: A comprehensive spelling checker for sinhala," *arXiv Prepr. arXiv:2107.02983*, 2021.
- [24] A. I. Fahma, I. Cholissodin, and R. S. Perdana, "Identification of Typographical Errors in Indonesian Language Documents Using N-Gram and Levenshtein Distance Methods," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 1, pp. 53–62, 2018.
- [25] B. McDonnell, "A conservative vowel phoneme inventory of sumatra: The case of besemah," *Ocean. Linguist.*, vol. 47, no. 2, pp. 409–432, 2009, doi: 10.1353/ol.0.0020.
- [26] N. Kharisma, N. Nadra, and R. Reniwati, "Phonology of Minangkabau Language Isolect Sikucur," *Diglosia J. Kaji. Language, Literature and Its Teaching*, vol. 4, no. 4, pp. 425–440, 2021, doi: 10.30872/diglosia.v4i4.280.
- [27] C. R. Fortin and D. Brodtkin, "Minangkabau -i: A locative, transitivizing, iterative, adversative suffix," *Proc. Linguist. Soc. Am.*, vol. 2, p. 42, 2017, doi: 10.3765/plsa.v2i0.4098.
- [28] F. S. Jumeilah, "Implementation of Support Vector Machine (SVM) for Research Categorization," *J. RESTI (System Engineering and Technol. Information)*, vol. 1, no. 1, pp. 19–25, 2017, doi: 10.29207/resti.v1i1.11.
- [29] D. Alita and A. R. Isnain, "Sarcasm Detection in Sentiment Analysis Process Using Random Forest Classifier," *J. Komputasi*, vol. 8, no. 2, pp. 50–58, 2020, doi: 10.23960/komputasi.v8i2.2615.
- [30] A. P. Wibawa, F. A. Dwiyanto, I. A. E. Zaeni, R. K. Nurrohman, and A. Afandi, "Stemming javanese affix words using nazief and adriani modifications," *J. Inform.*, vol. 14, no. 1, p. 36, 2020, doi: 10.26555/jifo.v14i1.a17106.



Kampus
Merdeka
INDONESIA JAYA



IEEE
INDONESIA SECTION

CERTIFICATE

OF APPRECIATION AS AN AUTHOR

INTERNATIONAL SYMPOSIUM ON INFORMATION TECHNOLOGY AND DIGITAL INNOVATION (ISITDI) 2022
Padang, 27-28 July 2022

YUHANDRI

For The Paper Entitled "**Development of the Minangkabau Local Language Translation Machine Based on Stemming**"

DR. ENG. BUDI RAHMADYA
CHAIRMAN OF ISITDI

