# Bangla Spell Checker

Sanjida samanta
sanjidasamanta277@gmail.com

June 17, 2023

# Contents

# 1 Introduction

A software application or program feature that identifies Bangla misspelled words and notifies the user,also has it's own keyboard. Depending on the Bangla spell checker, the feature may automatically show the misspelled word or provide the user with a list of possible corrections. It gives alternative/correct spellings to the words you might have confused and misspelled. It scans each and every finds the words in it and compares each word with a well-known list of spelled words (dictionary). To find out the misspelled Bengali words use some algorithm like String Matching, Edit Distance which are work of intelligent mixed up. After implement, then its will be get the most probable suggestion for our misspelled Bengali words. If a file containing text, is given in this software, one can easily catch the mistakes and can correct them. For knowing about Bangla spelling, this software is useful.

# 2 Target Customers

1. Student
2. Teacher
3. Journalist who works in Bangla newspaper
4. Content Writer
5. Government officers
6. Who writes document in Bangla

# 3 Features & Description

In my proposed software,
**a) File Read and Detection:** The user can input any doc file. If any misspellings are found, then One window open and show the possible correct word for those misspelled word.
**b) Show Misspelled word:** The words of the file will be checked and misspelled or wrong words will be marked by red underline.

**c) Correcting Misspelled word with drop down box:** By clicking on the wrong or misspelled words, as suggestion, a bunch of correct and similar words will be shown.
**d)** If users want, they can add that word to my dictionary.

# 4 Algorithm for spell checker

There are several types of spell checker algorithms. On this thesis described about two algorithms:

1. **String Matching**
2. **Edit Distance**
3. **Levenshtein distance**

## 4.1 String Matching Algorithm

String matching is an algorithm that attempts to locate one or more strings within a given text. It is an important class of string algorithms that searches a given position for a single string or multiple strings.

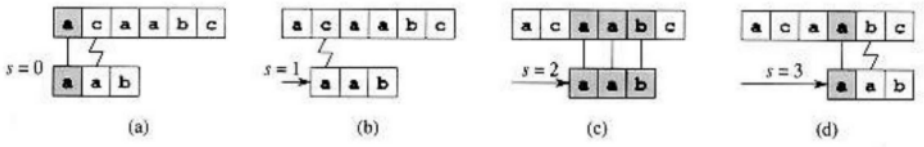


Figure 1: Processing on String Matching Algorithm

## 4.2 Edit Distance

Edit distance is a category of such type algorithm which determine the dissimilarity between two or multiple strings. There are numerous varieties of edit distance. they are –

1. Levenshtein distance
2. Longest common subsequence (LCS)

### 4.3 Levenshtein distance

This distance measures how much minimum operations or single character edits are needed to change one string to another. Single character edits mean insertion, substitution or deletion

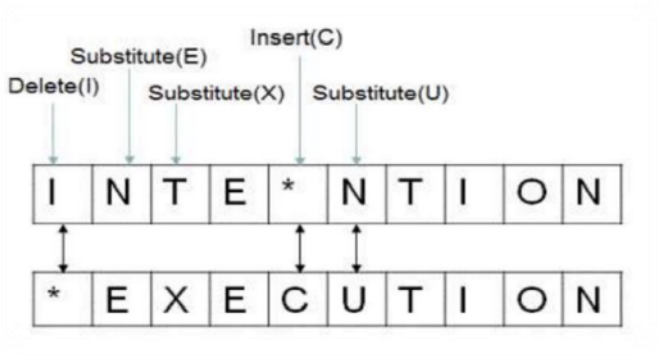


Figure 2: Processing on Levenshtein Distance

## 5 User Guide

**Step-1:** (Version Guideline)
- User must have java 17.0.2 or above version updated on their device.
**Step-2:** Go to link. JDK Download for Java download JDK 17.0.2:

1. Accept License Agreement
2. Run the exe for install

**Step-3:** Once you install Java in windows, click Close
**Step-4:** After installing JDK, check with any command prompt to make sure jdk is set in device environment
**Step-5:** Then download from here this project.
Bangla Spell Checker
**Step-6:**Then change Text file encoding in


Figure 3: JDK Installation

`UTF_8.`

$Properties \rightarrow Resource \rightarrow Text - file$ encoding in

`UTF_8.`

**Step-7:**After running my code, user will see such kind of Window
**Step-8:** When run the main method, user will see such kind of screen, Then after completing the loading Progress bar there will be open the below figure. In this part,user write some words and click check spell now? If the word is correct then it will remain the same.If not then there will be open the below figure
Here which word is not found that is show in not found in Dictionary. Here a list of possible words is shown in a table. If any word is click then show the Change to textField if change Button is click then change the word in spell checking window,if press ignore then ignore this one and if close then close the window.

- It also open files
- It also save files
- For Exit File
- For Change Theme


Figure 4: JDK Installation

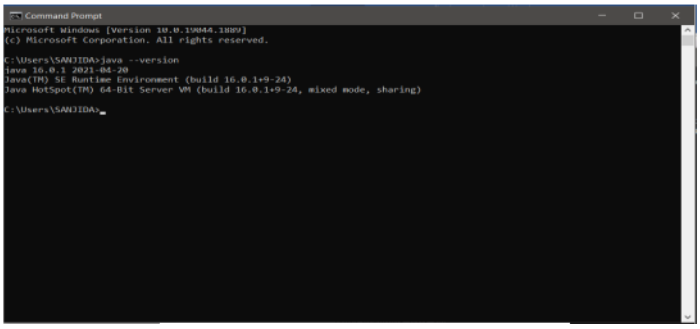


Figure 5: Java Version Check

# 6 Challenges

1. **Bangla word:** As unicodes of Bangla words are not well funrnished That's why it's a hard to handle whose unicodes.

2. **Bangla "Juktoborno":** As I worked with Bangla language, the uniqueness and difficulty also in the Bangla is juktoborno.

3. **File handling:** As I have restriction of using database, so I have to work with text file. And for storing huge amount of words and working with file was a challenge!

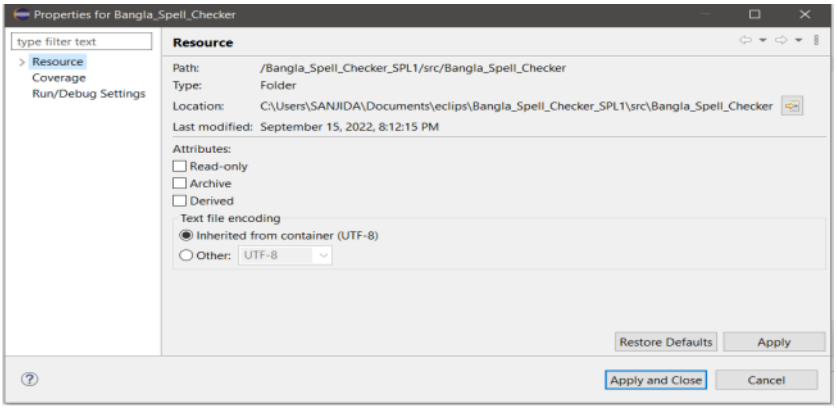


Figure 6: Encoding change

4. **Algorithm Implementation:** The algorithm I used it could create feature creep problem. The feature might not be handy and seems complex to the user.

5. **Tree and Recursion:** It was hard to create possible words using Tree and Recursion for Bangla as Bangla Unicode Doesn't well funrnished.

# 7 Future work

I will do more with spell checker. I will work with word suggestion and replace correct word with wrong word.It will work fine if we write the wrong word of a complex word like some Juktoborno and large length of word and using NLP concept.

# 8 Conclusion

In conclusion, while there are currently several options for Bangla spell checkers, there is room for improvement in terms of their accuracy and ability to handle context-dependent spelling errors. By exploring advanced techniques to keep up with changes in the language, it is possible to develop more effective and reliable spell checking tools that meet the needs of Bangla speakers.

# 9 Reference

[**1**] Introduction to Algorithms by Thomas H. Carmen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. [Third Edition]
[**2**] Java The Complete Reference by Herbert Scheldt. [ Eleventh Edition]
[**3**] Data Structures by SEYMOUR LIPSCHUTZ [Third Edition]


Figure 7: Starting Window

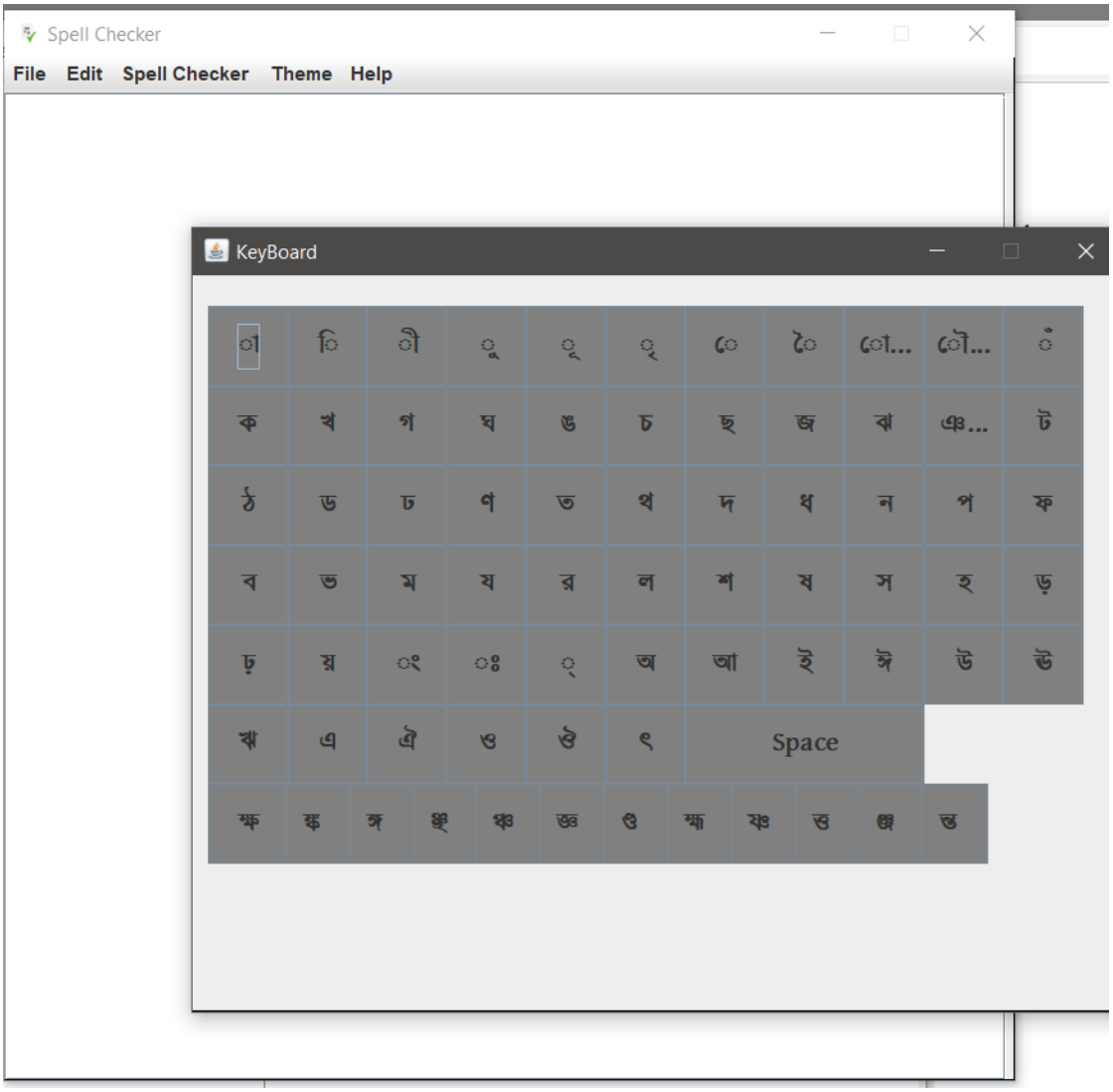


Figure 8: First Window

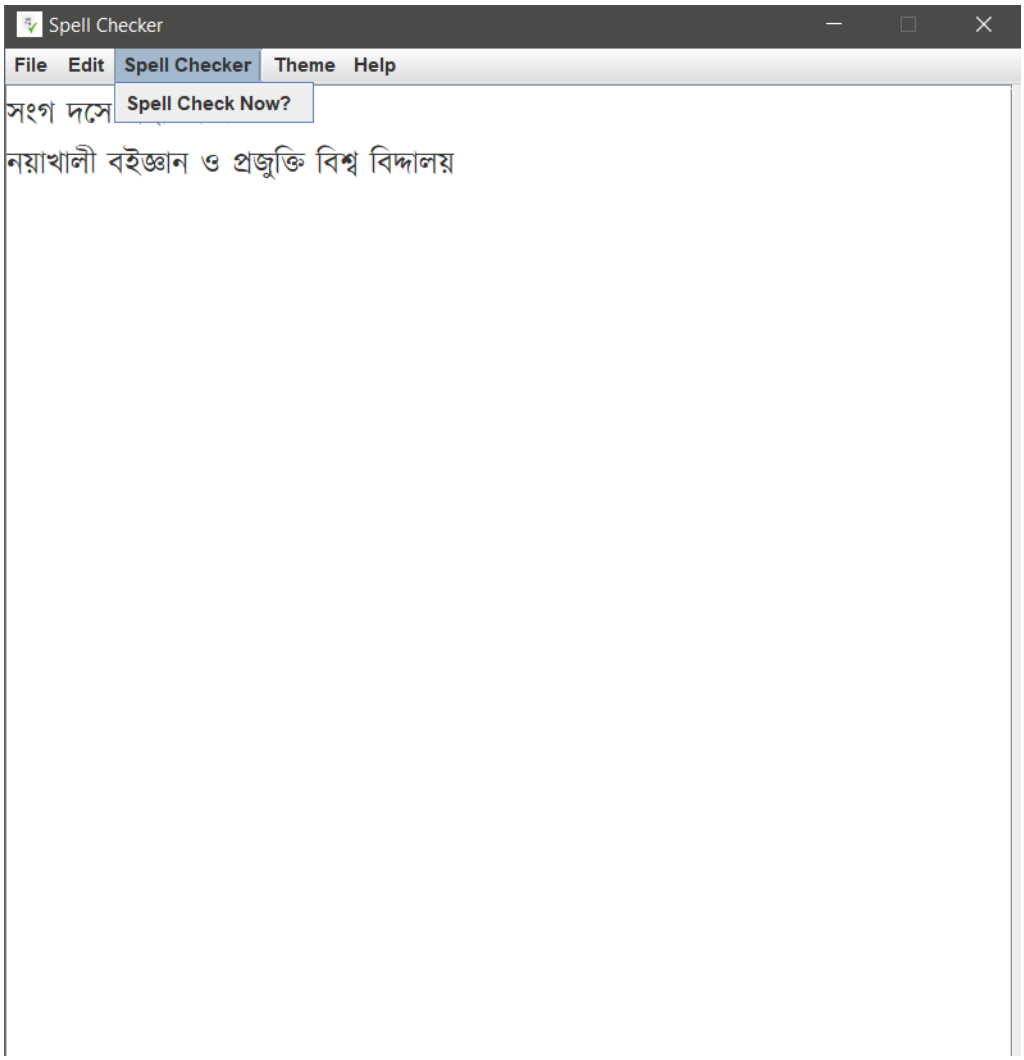


Figure 9: Spell Checking Window

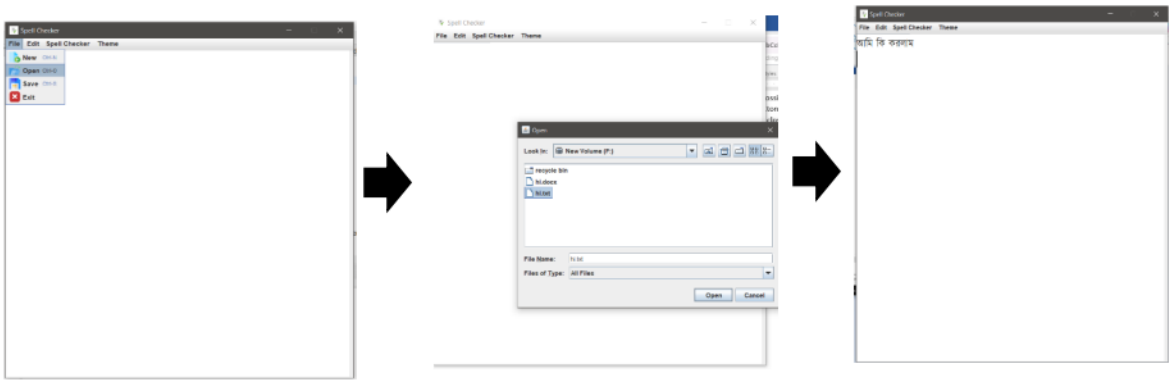

Figure 10: Open a File from Spell Checking Window

Figure 11: Save a File to Spell Checking Window

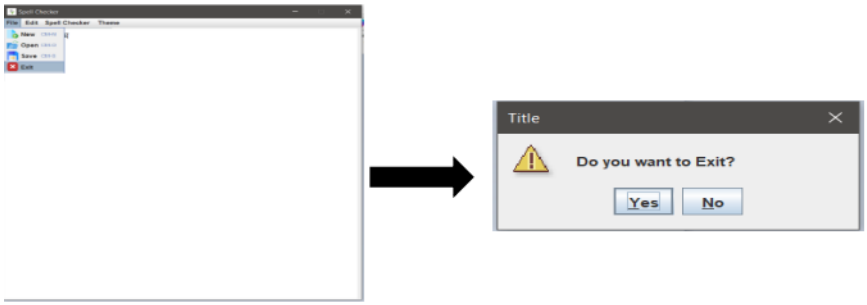


Figure 12: Exit File

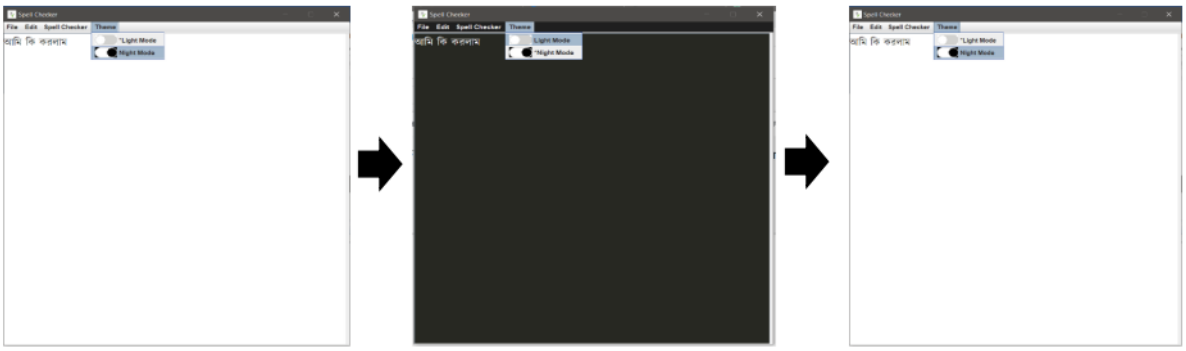

Figure 13: Theme of Spell Checking Window