THINK INDIA JOURNAL

ISSN:0971-1260

Vol-22-Issue-3-July-September-2019

Spell Checking Techniques- A Survey

Jagdish Kaur

ASSISTANT PROFESSOR

DEPARTMENT OF COMPUTER SCIENCE

DAV COLLEGE FOR WOMEN, FEROZEPUR

ABSTRACT-Spell checking is the process of detecting and correcting spelling errors and provides the best possible suggestions for spelling errors. This survey paper covers all the spelling detection and correction techniques.

Keywords- Error detection techniques, Error correction techniques.

I. INTRODUCTION

Spell Checker takes input from file and check if the word is present in the dictionary, if it is not present then it is considered as incorrect word otherwise it is correct word. It provides the best possible suggestions for incorrect words.

II. SPELLING ERRORS

There are two types of spelling errors: Real -word errors and non-word errors. Real words errors are those errors which are known words in dictionary but mistyped by mistake. Non-words are not known words in that cannot be found in the dictionary.

III. SPELL CHECKING DETECTION TECHNIQUES

A. Dictionary Lookup Technique

In this technique, every word of given input file is checked for its presence in the dictionary. The word is considered as incorrect word if it is not in the dictionary otherwise it is correct word. Various handcrafted rules are applied to generate the correct spelling suggestions for incorrect words.

B. N-gram based Technique

ISSN:0971-1260
Vol-22-Issue-3-July-September-2019

N-gram are n-letters of sequence of strings where n is one, two or three. One letter n-gram are known as unigram, two letters n gram as bigram and three letters referred to as a trigram. N-gram detection technique work by examining each n gram is an input string and looking it up in a precompiled table of n – gram statistic to as certain either its existence or its frequency of words or strings that are found to contain non-existence or highly infrequent n-grams are identified as either misspellings [1].

IV. SPELL CHECKING CORRECTION TECHNIQUES

A. Rule based Technique

In this approach handcrafted rules are made correspondence to the features of Punjabi Language. These rules are applied on that words which are not present in the dictionary. These rules help to transform misspelling words into valid words by generating correct spelling suggestions.

B. Edit Distance Technique

In this technique minimum edit distance is computed by comparing misspelled word with every word of dictionary. After comparison word with minimum edit distance are selected as correct alternatives suggestion.

C. Probabilistic technique

Two common methods are available as transition probabilities and confusion probabilities. In transition probability a given letter is followed by another given letter. This is not useful when data is accessed from dictionary. Confusion probabilities estimate how a given character is mistaken or substituted for another given letter.

D. N-gram based Technique

N-gram can be used with dictionary or without dictionary. Using without dictionary, n-gram are used to find in which position misspelled word occur. Used with dictionary, n-grams are used to find the distance between words and the words are always checked against the dictionary.

E. Neural Network

Neural Network mostly used back propagation algorithm for training it.

ISSN:0971-1260
Vol-22-Issue-3-July-September-2019

Back propagation net consists of three layers: input layer, hidden layer and output layer. Every node in input layer is connected to every node in hidden layer by a weighted link and each node in hidden layer is connected by weighted link to every node in output layer.

V. CONCLUSION

This paper presented a study on various spelling error detection and correction Techniques. Types of spelling errors also discussed in this. A lot of work has been done

For English but there is not much working done in the area of Punjabi language . So, in

Future I am looking for design and implement spell checker for Punjabi language.

REFERENCES

[1] R. Mishra and N.Kaur," A Survey of spelling error detection and correction techniques", International Journal of Computer Trends and Technology, vol.41, Issue.3,2013.

[2] G.S Lehal and M.Bhagat," Spelling Error Pattern Analysis of Punjabi Typed Text", In Proceeding of International Symposum on Machine Translation NLP and TSS,pp.128-141.

[3] N .Gupta and P.Mathur,"Spell Checking techniques in NLP:A survey", International Journal of Advanced Research in Computer Science and Software Engineering,vol.2, Issue.12, December 2012.

[4] G.S Lehal,"Design and Implementation of Punjabi Spell Checker",International Journal of Systemics,Cybemeticsand Infomatics.

[5] B.Kaur," Review on Error Detection and Error Correction Techniques in NLP:",international Journal of advanced Research in Computer Science and Software Engineering,vol.4,Issue.6,2014.`1