*Nermeen Khalid*

*Yasmeen Ahmed*

*Abdallah Gamal*

*Mohanad Ehab*

*Ali Mohamed*

NLP

Spelling Correction and Detection Report

5/1/2018

Helwan Unisveristy

Table of Contents

[Abstract : 3](#_Toc512957487)

[Data Set : 3](#_Toc512957488)

[Source : https://github.com/dwyl/english-words/blob/master/words.txt 3](#_Toc512957489)

[Methodology: 3](#_Toc512957490)

[Experiment 1: 3](#_Toc512957491)

[Execution time: 3](#_Toc512957492)

[Experiment 2: 3](#_Toc512957493)

[Execution time: 4](#_Toc512957494)

[Experiment 3: 4](#_Toc512957495)

[Execution time: 4](#_Toc512957496)

[Results: 4](#_Toc512957497)

[Discussion: 4](#_Toc512957498)

[Conclusion: 5](#_Toc512957499)

# Abstract :

Spelling correction and detection project aims at detecting user’s spelling mistakes and correcting them ,By using the Edit Distance algorithm , which compares two strings and finds the minimum edits done to change first word to the second , different algorithms for edit distance were implemented to compare performance measures . Different algorithms allow for transposition , substitution , deletion and insertion to transform a word to another .

# Data Set :

Data set used was a text file containing various words that were used to compare words entered by the user , and correcting them .

## Source : <https://github.com/dwyl/english-words/blob/master/words.txt>

# Methodology:

## Experiment 1:

In experiment one we used the Damerau edit distance method , where we allowed for transposition operation , a string word was used as the first argument for the method , and compared with the data set which is passed as the second argument to the method , all edit distance operations are done on the word to match it with the data set and return the correct word .

### Execution time:

## Experiment 2:

In experiment two we used regex to try to match a correct word from the data set , all text from the data set were changed to lower case , a list of known words was made by getting the top 10 words in the data set , 2 methods were then implemented each calculates different edit distance ex words that are one edit away from the entered word etc , then the correct word is returned.

### Execution time:

## Experiment 3:

### Execution time:

# Results:

# Discussion:

# Conclusion:

Spelling correction is an important field , used by many websites and application as humans are prone to make mistakes while writing , and to get best performance and user experience word detection and correction is used , edit distance algorithm is an efficient algorithm and is used by large search engines like Google . It allows for different configurations based on the application requirements, it’s a dynamic programming algorithm and hence its efficient for large applications .