# Minimum Edit Distance(MED) Algorithm

#### I. INTRODUCTION

Minimum edit distance (MED) or Levenshtein algorithm is used to rate the similarity between two sequences. In practice it is used to rate the similarity between words in search results.

In this study, a Levenshtein algorithm was implemented in Java. The program takes input as text and lists all words written incorrect with their first 5 alternative correct. It uses a text file as a sample word pool to perform these operations. The program outputs were analyzed separately according to the total running time and words - alternative correct words list for 20 words.

#### II. DEFINITION OF THE ALGORITHM

A. Methods, Inputs, and Outputs of the Program

The following methods were used in the program:

- distance method- MED algorithm was used in this method. The actual process. Comparing is done by this method. Distance method is hearth of the code [1].
- main method
  - ✓ Taking inputs,
  - ✓ File operations, [2]
  - ✓ Calculating total running time, [3]

are included in this method.

- Inputs- Any text.
- Outputs- Total running time and all words written incorrect with their first 5 alternative correct words.

#### III. SCREENSHOTS OF THE PROGRAM

The program screenshots are shown in the following.

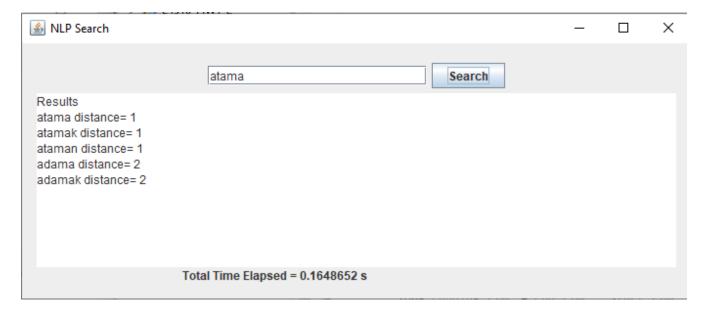


Figure 1: Outputs and total running time for the test word "atama".

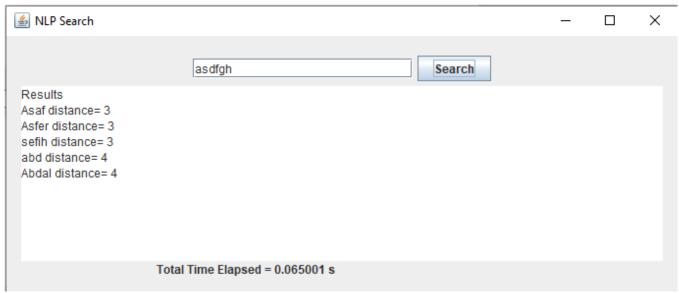


Figure 2: Outputs and total running time for a test word that generated randomly.

## IV. TEST RESULTS

**Table 1:** Test words, their alternatives and total running times for each word.

bozukluk distance= 2 çabukluk distance= 2 çabukluk distance= 2 çocukluk equiva distance= 2 çocukluk distance= 2 çocuklar distance= 2 çikak distance= 2 çikak distance= 2 çikak distance= 2 çikik distance= 2 çikik distance= 2 çikik distance= 2 çokekl distance= 2 çokek distance= 2 rabita distance= 0 rabit distance= 1 rabita zabita distance= 1 rabita distance= 1 rabit distance= 1 rabit distance= 1 rabit distance= 1 rabit distance= 1 raki distance= 4 radikallşmek distance= 4 radikallşmek distance= 4 radikallşmek distance= 4 radikallişmek distance= 2 meyusiyet distance= 2 meyusiyet distance= 2 meyusiyet mesuliyet distance= 2	Test words	Alternative words for test word	Total running time(s)
cocukluk cocukçu distance= 2 cocuklar distance= 1 cocumentaria cocumentar		bozukluk distance= 2	
cocuklar distance= 2 coğunluk distance= 2 çıkak distance= 2 çıkık distance= 2 çıkık distance= 2 çıkık distance= 2 çıkık distance= 2 rabta distance= 0 rabta distance= 1 rabıta zabıta distance= 1 rabıtalı distance= 2 rasıt distance= 2 rasıt distance= 1 rabı distance= 1 rabı distance= 1 rakı distance= 4 radıkallışmek distance= 2 meyusiyet distance= 2 meyusiyet distance= 2		çabukluk distance= 2	
çoğunluk distance= 2 çıkık distance= 2 çıkık distance= 2 çekik distance= 2 çekik distance= 2 çökek distance= 2 rabıta distance= 0 rabıt distance= 1 rabıta zabıta distance= 1 rabıtalı distance= 1 rabıt distance= 1 rakı distance= 1 rakı distance= 1 rakı distance= 1 araba distance= 1 araba distance= 4 dikleşmek distance= 4 radikallşmek dikleşmek distance= 4 radikallzm distance= 4 radikalizm distance= 4 radikalizm distance= 4 radikalizm distance= 2 meyusiyet mesuliyet distance= 2 0.00254	çocukluk	çocukçu distance= 2	0.0202
çıklak distance= 2 çıklık distance= 2 rabıta distance= 0 rabıta distance= 1 rabıta   zabıta distance= 1 rabıtalı distance= 2 rasıt distance= 1 rabıt distance= 1 rabıt distance= 1 rabıt distance= 1 rakı distance= 1 rakı distance= 1 rakı distance= 1 rakı distance= 4 radıkallışmek distance= 4 radikallışmek distance= 2 meyusiyet distance= 2 meyusiyet distance= 2		çocuklar distance= 2	
çıkık distance= 2 rabıta distance= 0 rabıt distance= 1 rabıta zabıta distance= 1 rabıtalı distance= 2 rabıtalı distance= 1 rabıt distance= 1 rakı distance= 1 rakı distance= 1 araba distance= 4 radikalleşmek distance= 4 radikalleşmek distance= 4 radikallizm distance= 2 meyusiyet distance= 2 meyusiyet mesuliyet distance= 2 0.00254		çoğunluk distance= 2	
ccklk cilk distance= 2 ckik distance= 2 ckik distance= 2 ckik distance= 2 ckik distance= 0 cabit distance= 0 cabit distance= 1 cabit distance= 1 cabit distance= 1 cabit distance= 2 cabit distance= 2 cabit distance= 2 cabit distance= 1 cabit distance= 2 cabit distance= 2 cabit distance= 1 cabit distance= 2 cabit distance= 1 cabit distance= 1 cabit distance= 1 cabit distance= 2 cabit distance= 1 cabit distance= 2 cabit dista		çıkak distance= 2	
çekik distance= 2 çökek distance= 0 rabıta distance= 1 rabıta distance= 1 rabıtalı distance= 2 rabıtalı distance= 2 rabıtalı distance= 2 rabıtalı distance= 2 rabıt distance= 1 rakı distance= 1 rakı distance= 1 araba distance= 2 radikalleşmek distance= 4 radikalleşmek distance= 4 radikalizm distance= 4 radikalizm distance= 4 radikalizm distance= 4 radikalleşmek distance= 2 meyusiyet distance= 2 meyusiyet distance= 2		çıkık distance= 2	
rabita distance= 2 rabita distance= 1 rabita distance= 1 rabita distance= 1 rabitali distance= 2 rabitali distance= 2 rabitali distance= 2 rabit distance= 1 raki distance= 1 raki distance= 1 araba distance= 1 araba distance= 2  radikalleşmek distance= 1 adileşmek distance= 4 radikalişmek distance= 4 radikalizm distance= 4 radikalleşmek distance= 3 adileşmek distance= 4 radikalleşmek distance= 4 radikalizm distance= 4 radikalizm distance= 2 meyusiyet distance= 2 meyusiyet mesuliyet distance= 2 0.0254	çcklk	cılk distance= 2	0.0163
rabita distance= 0 rabit distance= 1 zabita distance= 1 zabita distance= 2 rabitali distance= 2 rasit distance= 2 rabit distance= 1 rabit		çekik distance= 2	
rabita istance= 1 zabita distance= 1 rabitali distance= 2 rasit distance= 2 rasit distance= 1 rabit distance= 1 rabit distance= 1 rabit distance= 1 raki distance= 1 raki distance= 1 araba distance= 2  radikalleşmek distance= 1 adileşmek distance= 4 radikallişmek dikleşmek distance= 4 radikalleşmek distance= 2 meyusiyet distance= 2 meyusiyet mesuliyet distance= 2 0.00254		çökek distance= 2	
rabita zabita distance= 1		rabita distance= 0	
rabitali distance= 2 rasit distance= 1 rabit distance= 1 rabit distance= 1 raki distance= 1 raki distance= 1 araba distance= 2  radikalleşmek distance= 1 adileşmek distance= 4 radikallşmek dikleşmek distance= 4 radikalizm distance= 4 radikalizm distance= 4 radikalleşmek distance= 4 radikalizm distance= 4 radikalleşmek distance= 2 meyusiyet distance= 2 meyusiyet distance= 2 meyusiyet mesuliyet distance= 2		rabit distance= 1	
rasit distance= 2  rab distance= 1 rabit distance= 1 rabit distance= 1 raki distance= 1 raki distance= 1 araba distance= 2  radikalleşmek distance= 1 adileşmek distance= 4 dikleşmek distance= 4 radikalizm distance= 4 radikalizm distance= 4 radikalleşmek distance= 2 meyusiyet distance= 2	rabıta	zabita distance= 1	0.0169
rabit distance= 1 rabit distance= 1 raki distance= 1 raki distance= 1 araba distance= 2  radikalleşmek distance= 1 adileşmek distance= 4 dikleşmek distance= 4 radikallişmek  radikalleşmek distance= 4 radikallişmek distance= 4 radikal distance= 4 radikal distance= 4 radikalizm distance= 4 radikalizm distance= 4 radikalizm distance= 2 meyusiyet mesuliyet distance= 2 0.0254		rabitali distance= 2	
rabit distance= 1 raki distance= 1 raki distance= 1 raki distance= 1 araba distance= 2  radikalleşmek distance= 1 adileşmek distance= 4 radikallşmek dikleşmek distance= 4 radikalizm distance= 4 radikalleşmek distance= 3 adileşmek distance= 4 radikalleşmek distance= 4 radikalleşmek distance= 4 radikallizm distance= 2 meyusiyet distance= 2 meyusiyet mesuliyet distance= 2 0.00254		rasıt distance= 2	
rakı distance= 1 rakı distance= 1 rakı distance= 2  radikalleşmek distance= 1 adileşmek distance= 4 radikallşmek  dikleşmek distance= 4 radikalizm distance= 4 radikalleşmek distance= 3 adileşmek distance= 3 adileşmek distance= 4 radikalleşmek distance= 4 radikalleşmek distance= 4 meyusiyet distance= 4 meyusiyet distance= 0 memuriyet distance= 2 meyusiyet  no.0154  0.0154		rab distance= 1	
rakı distance= 1 araba distance= 2  radikalleşmek distance= 1 adileşmek distance= 4 radikallşmek  dikleşmek distance= 4 radikalizm distance= 4 radikalizm distance= 3 adileşmek distance= 4 radikleşmek distance= 4 radikleşmek distance= 4 radikleşmek distance= 4 meyusiyet distance= 4 meyusiyet distance= 0 memuriyet distance= 2 meyusiyet mesuliyet distance= 2 0.0254		rabit distance= 1	
radikalleşmek distance= 1 adileşmek distance= 4 radikallşmek dikleşmek distance= 4 radikalizm distance= 4 radikalleşmek distance= 3 adileşmek distance= 4 radikalleşmek distance= 4 radikalleşmek distance= 4 radikllşmk dikleşmek distance= 4 radikal distance= 4 radikalizm distance= 4 radikalizm distance= 4 meyusiyet distance= 0 memuriyet distance= 2 meyusiyet mesuliyet distance= 2 0.0254	rabı	rakı distance= 1	0.0154
radikalleşmek distance= 1 adileşmek distance= 4 dikleşmek distance= 4 radikallşmek  radikalizm distance= 4 radikalizm distance= 3 adileşmek distance= 4 radikllşmek distance= 4 radikllşmek distance= 4 radikllşmek distance= 4 radikal distance= 4 radikal distance= 4 radikalizm distance= 2 meyusiyet distance= 2 meyusiyet distance= 2 meyusiyet distance= 2 0.0254		rakı distance= 1	
adileşmek distance= 4 dikleşmek distance= 4 radikalizm distance= 4 radikalizm distance= 3 adileşmek distance= 3 adileşmek distance= 4 radikllşmk dikleşmek distance= 4 radikal distance= 4 radikal distance= 4 radikalizm distance= 4 radikalizm distance= 4 meyusiyet distance= 0 memuriyet distance= 2 meyusiyet mesuliyet distance= 2 0.0254		araba distance= 2	
radikallşmek distance= 4		radikalleşmek distance= 1	
kadifeleşmek distance= 4 radikalizm distance= 4 radikalleşmek distance= 3 adileşmek distance= 4 radikllşmk dikleşmek distance= 4 radikal distance= 4 radikalizm distance= 4 meyusiyet distance= 0 memuriyet distance= 2 meyusiyet mesuliyet distance= 2 0.0254	radikallşmek	adileşmek distance= 4	
radikalizm distance= 4  radikalleşmek distance= 3 adileşmek distance= 4  radikllşmk  dikleşmek distance= 4  radikal distance= 4  radikalizm distance= 4  meyusiyet distance= 0  memuriyet distance= 2  meyusiyet  mesuliyet distance= 2  0.0254		dikleşmek distance= 4	0.0225
radikalleşmek distance= 3 adileşmek distance= 4 radikllşmk  dikleşmek distance= 4 radikal distance= 4 radikalizm distance= 4 meyusiyet distance= 0 memuriyet distance= 2 meyusiyet mesuliyet distance= 2 0.0254		kadifeleşmek distance= 4	
radikllşmk distance= 4  dikleşmek distance= 4  radikal distance= 4  radikalizm distance= 4  meyusiyet distance= 0  memuriyet distance= 2  meyusiyet distance= 2  0.0254		radikalizm distance= 4	
radikllşmk distance= 4 0.0209 radikal distance= 4 radikalizm distance= 4 meyusiyet distance= 0 memuriyet distance= 2 meyusiyet mesuliyet distance= 2 0.0254	radikllşmk	radikalleşmek distance= 3	
radikal distance= 4 radikalizm distance= 4 meyusiyet distance= 0 memuriyet distance= 2 meyusiyet mesuliyet distance= 2 0.0254		adileşmek distance= 4	
radikalizm distance= 4  meyusiyet distance= 0 memuriyet distance= 2 meyusiyet mesuliyet distance= 2 0.0254		dikleşmek distance= 4	0.0209
meyusiyet distance= 0 memuriyet distance= 2 meyusiyet mesuliyet distance= 2 0.0254		radikal distance= 4	
memuriyet distance= 2 meyusiyet mesuliyet distance= 2 0.0254		radikalizm distance= 4	
meyusiyet mesuliyet distance= 2 0.0254		meyusiyet distance= 0	
		memuriyet distance= 2	
mozunivot diatonos- 2	meyusiyet	mesuliyet distance= 2	0.0254
mezumyet distance= 2		mezuniyet distance= 2	

	haysiyet distance= 3	
	meyusiyet distance= 3	
	maksi etek distance= 4	
myusiytek	mersiye distance= 4	0.0200
	meyus etmek distance= 4	
	musibet distance= 4	
	mezcetmek distance= 0	
	bezetmek distance= 2	
mezcetmek	cezbetmek distance= 2	0.0218
	dercetmek distance= 2	
	menetmek distance= 2	
	mezcetmek distance= 1	
	bezetmek distance= 2	
mezctmek	menetmek distance= 2	0.0237
	mercimek distance= 2	
	ahzetmek distance= 3	
	mıhlamak distance= 0	
	mıhlanmak distance= 1	
mıhlamak	çınlamak distance= 2	0.0335
	cızlamak distance= 2	
	dışlamak distance= 2	
	ahlak distance= 2	
	çıplak distance= 2	
mıhlak	ıtlak distance= 2	0.0176
	dılak distance= 2	
	fırlak distance= 2	
istekle	istekle distance= 0	
	istekli distance= 1	
	iskemle distance= 2	0.0256
	istek distance= 2	
	isteka distance= 2	
	içtenlikle distance= 3	
	istekle distance= 3	
isteklekle	isteklenmek distance= 3	0.0216
	içtenlikle distance= 3	
	destekleme distance= 4	
	destruine distance— T	

çocuk bakıcısı	çocuk bakıcısı distance= 0	
	çocuk bakıcı distance= 2	
	at bakıcısı distance= 5	0.0316
	at bakıcısı distance= 5	
	çocuk yazını distance= 5	
	çokça distance= 3	
	acaba distance= 4	
çock ba	çıkma distance= 4	0.0170
	caba distance= 4	
	çaba distance= 4	
	çocukçu distance= 1	
çocukça	çocukçu distance= 1	
	çabukça distance= 2	0.0172
	çocuk distance= 2	
	çocuklar distance= 2	
	çokça distance= 1	
	çabukça distance= 2	
çoukça	çocukçu distance= 2	0.0185
	çokçu distance= 2	
	çocukçu distance= 2	
	fakir distance= 0	
fakir	bakir distance= 1	
	fakr distance= 1	0.0166
	fikir distance= 1	
	hakir distance= 1	
fakitl	akil distance= 2	
	akit distance= 2	
	fail distance= 2	0.0177
	fakat distance= 2	
	fakir distance= 2	

As shown in the Table 1, the program tested with 20 different words. The running times were computed for each word in seconds. A word was written correctly. Then, random changes were made in the letters and results were reached. These steps were performed for 10 different words. A total of 20 test words, alternative lists and running times were obtained. At the end of the study, it was found that with the same order, the similarity of letters increased and the distance and running times decreased.

### REFERENCES

- [1] https://rosettacode.org/wiki/Levenshtein\_distance#Java
- $[2] \quad https://stackoverflow.com/questions/1844688/how-to-read-all-files-in-a-folder-from-java$
- [3] https://github.com/belenyasin/Java-Gecen-Sure-Hesaplama