CHAPTER 5
IMPLEMENTATION AND TESTING

5.1 Implementation

$$d_j = \begin{cases} 0 & \text{if } m = 0 \\ \frac{1}{3}\left(\frac{m}{|s_1|} + \frac{m}{|s_2|} + \frac{m-t}{m}\right) & \text{otherwise} \end{cases} \quad (1)$$

$$\left\lfloor \frac{\max(|s_1|, |s_2|)}{2} \right\rfloor - 1 \quad (2)$$

$$d_w = d_j + \left(l \times p\left(1 - d_j\right)\right) \quad (3)$$

The following is an example of calculating the Jaro Winkler algorithm. If string 1 is pomap and string is 2 pumps, then:

m = 5

s1 = 5

s2 = 5

the only swapped characters are P and A, so the number of transpositions (t) = 1.

then the jaro distance :

dj = $\frac{1}{3}\left(\frac{5}{5} + \frac{5}{5} + \frac{4}{5}\right)$ = 0.933

then from the arrangement of s1 and s2 it can be seen that the same character length until found inequality is P, O, M so that the value of l = 3 with a constant value p = 0.1. then the jaro distance:

dw = 0.933+(3x0.1(1-0.933)) = 0.953

So, the jaro distance between POMAP and POMPA is 0.953

Levensthein distance :

d(i-1,j)+1

d(i,j-1)+1

d(i-1,j-1)+cost

Table 5.1: Levensthein distance between POMAP and POMPA

		P	O	M	A	P
	0	1	2	3	4	5
P	1	0	1	2	3	4
O	2	1	0	1	2	3
M	3	2	1	0	1	2
P	4	3	2	1	0	1
A	5	4	3	2	1	1

5.2 Testing

There are 50 sentence data that will be tested in this research. Each sentence consists of 3 words. Each words will be calculated the distance with 3 word suggestions from Kamus Besar Bahasa Indonesia with cases of addition, deletion, replacement, and exchange of letters.