Report based on findings on SnowballStemmer

(10/01/2019)

–Ankush Chaudhari, 17BCE0778

–Harshil Shah, 17BCE2411

The following words were tested on the SnowballStemmer to find out the inconsistencies in the results of the stemming algorithm and to figure out some of its rules:

Input	Output Stem	Comment
Kiting	Kite	Simple input, output ends with 'e'
Biting	Bite	Simple input, output ends with 'e'
Riding	Ride	Simple input, output ends with 'e'
Spiting	Spite	Simple input, output ends with 'e'
Giving	Give	Simple input, output ends with 'e'
Mining	Mine	Simple input, output ends with 'e'
Fighting	Fight	Two or more consonants just preceding the 'ing', output does not end with 'e'
Pitting	Pit	Two or more consonants just preceding the 'ing', output does not end with 'e'
Riddling	Riddl	Two or more consonants just preceding the 'ing', output does not end with 'e'
Spitting	Spit	Two or more consonants just preceding the 'ing', output does not end with 'e'
Manning	Man	Two or more consonants just preceding the 'ing', output does not end with 'e'
Igniting	Ignite	Two or more vowels in the part of the word before the 'ing', output does not end with an 'e'
Criticising	Criticis	Two or more vowels in the part of the word before the 'ing', output does not end with an 'e'
Exiting	Exit	Two or more vowels in the part of the word before the 'ing', output does not end with an 'e'
Accepting	Accept	Two or more vowels in the part of the word before the 'ing', output does not end with an 'e'
Seeking	Seek	Two or more vowels in the part of the word before the 'ing', output does not end with an 'e'
Eviting	Evit	Two or more vowels in the part of the word before the 'ing', output does not end with an 'e'
Imagining	Imagin	Two or more vowels in the part of the word before the 'ing', output does not end with an 'e'
Carrying	Carri	'Y' preceding the 'ing' part of the word turns into 'i'
Marrying	Marri	'Y' preceding the 'ing' part of the word turns into 'i'

Based on the words tested above, the following 3 rules about the Snowball Stemmer can be established upon words with an 'ing' suffix:

1) When in a word with an 'ing' suffix has a maximum of only one vowel in the part of the word before the 'ing' and a maximum of one consonant immediately preceding the 'ing', the stem of the word ends with an 'e'.
2) When the word does not follow any 1 or both of the above mentioned conditions, the stem of the input word does not end with an 'e', often resulting in incorrect results.
3) When the word has the letter 'y' immediately preceding the 'ing', the stem of the word has the original word with the 'ying' replaced by an 'i'.