Assignment - 1

Subject: **Natural langusge Processing**

Sr. No.	Word Passed	POS	Stem Word Obtained	Actual Stem Word	Is the result correct?
1.	boys	Noun	boy	Boy	Yes
2.	boy	Noun	Boy	Boy	Yes
3.	toys	Noun	Toy	Toy	Yes
4.	toy	Noun	Toy	Toy	Yes
5.	Noises	Noun	Nois	Noise	No
6.	Noise	Noun, verb	Nois	Noise	No
7.	noised	verb	Nois	Noise	No
8.	Noising	verb	Nois	Noise	No
9.	Voices	Noun	Voic	Voice	No
10.	Voice	Noun, verb	Voic	Voice	No
11.	voicing	verb	voic	Voice	No
12.	voiced	verb	voic	Voice	No
13.	cry	verb	cri	cry	No
14.	cried	verb	cri	cry	No
15.	cries	verb	cri	cry	No
16.	try	verb	tri	try	No
17.	tried	verb	tri	try	No
18.	tries	verb	tri	try	No
19.	arrive	verb	arriv	arriv	yes
20.	arrival	noun	arriv	arriv	yes
21.	arrived	verb	arriv	arriv	yes
22.	arriving	verb	arriv	arriv	yes
23.	arriver	noun	arriv	arriv	yes
24.	said	verb	said	say	No
25.	say	verb	say	say	Yes
26.	heard	verb	heard	hear	No
27.	hear	verb	hear	hear	Yes
28.	heart	noun	heart	heart	Yes
29.	survive	verb	surviv	surviv	Yes
30.	survived	verb	surviv	surviv	Yes

31.	surviving	verb	surviv	surviv	Yes
32.	survival	noun	surviv	surviv	Yes
33.	survivor	noun	survivor	survivor	Yes
34.	trivial	noun	trivial	trivial	yes
35.	Semantics	noun	semant	semantics	no
36.	analysis	noun	analysi	analyses	no
37.	division	noun	divis	divide	no
38.	owing	verb	owe	owe	yes
39.	owner	noun	owner	own	no
40.	seeing	verb	see	see	yes
41.	logitudinal	noun	longitudin	longitude	no
42.	latitudinal	noun	latitudin	latitude	no
43.	mare	noun	mare	mare	yes
44.	children	noun	children	child	no
45.	helper	noun	helper	helper	yes
46.	lamented	verb	lament	lament	yes
47.	lamenting	verb	lament	lament	yes
48.	am	verb	am	be	no
49.	were	verb	were	be	no
50.	being	verb	be	be	yes
51.	I	pronoun	I	I	yes
52.	me	pronoun	me	I	no
53.	paste	verb	past	paste	no
54.	his	pronoun	hi	his	no
55.	was	verb	wa	be	no
56.	has	verb	ha	be	no
57.	gas	noun	ga	gas	no
58.	face	noun	face	face	yes
59.	faces	noun	face	face	yes
60.	less	noun	less	less	yes
61.	loss	noun	loss	loss	yes
62.	lost	verb	lost	lose	no
63.	lose	verb	lose	lose	yes
64.	her	pronoun	her	her	yes
65.	sing	verb	yes	yes	yes
66.	flying	verb	fli	fly	no
67.	xerox	noun	xerox	xerox	yes

68.	politics	noun	polit	politics	no
69.	polite	adverb	polit	polite	no
70.	flight	verb	flight	fly	no
71.	pros	noun	pro	pros	no
72.	cons	noun	con	cons	no
73.	mess	noun	mess	mess	yes
74.	maze	noun	mazze	maze	yes
75.	is	verb	is	be	no
76.	does	verb	doe	do	no
77.	aid	noun	aid	aid	yes
78.	made	verb	made	make	no
79.	make	verb	make	make	yes
80.	might	verb	might	may	no
81.	May	verb	may	may	yes
82.	could	verb	could	can	no
83.	can	verb	Can	can	yes
84.	would	verb	would	will	no
85.	will	verb	will	will	yes

Proposed rules (Modification in Porter Stemming Algorithm) :

1. For cried, tried, ===>
Modify/ add rule in 1(b), (*v*)ed ->
          (*r)ied -> y

2. For cries, tries etc ==>
Modify/ add rule in 1(a) ies -> i
(*r)ies -> y

3. For said etc, ===>
(*v)id -> y

4. For noises, voices ===>
add rule in 1(a)
Condition-> (*i_)es -> e

5. For his, gas,
Modify/ add rule in 1(a)
(m=0) and (*v)s -> s
i.e.it contain a vowel before s, then don't convert the word

6. For logitudinal, latitudinal, etc. ===>
Modify/ add rule in 2
(m>=1) and (*d)inal -> e