Nistha Adhikari

Text Classification

Classifying the text using bag of words approach:

Text	Label
Congrats, You have won!! reply to our sms for a free nokia mobile + free camcorder.	spam
Congrats! 1 year special cinema pass for 2 is yours. reply to this sms to claim your prize.	spam
I am pleased to tell you that you are awarded with a 1500 Bonus Prize, reply to this sms to claim your prize.	spam
Dont worry. I guess he is busy.	not spam
Going for dinner. msg you later.	not spam
Ok, I will call you up when I get some cash.	not spam

Unique words ignoring case and punctuation

Vocabulary

Congrats, you, have, won, reply, to, our, sms, for, a, free, nokia, mobile, +, camcorder, 1, year, special, cinema, pass, for, 2, is, yours, this, claim, your, prize, I, am, pleased, tell, that, are, awarded, with, 1500, bonus, don't, worry, guess, he, busy, going, dinner, msg, later, ok, call, up, when, get, some, cash

	Congrats	you	have	won	repl y	to	ou r	sms	for	a	free	nokia	mobile	+	camcorder	1	year	special	cinema	pass
Congrats,You have won!!reply to our sms for a free nokia mobile + free camcorder.	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1	0	0	0	0	0
Congrats! 1 year special cinema pass for 2 is yours. reply to this sms to claim your prize.	1	0	0	0	1	2	0	1	1	0	0	0	0	0	0	1	1	1	1	1
I am pleased to tell you that you are awarded with a 1500 Bonus Prize, reply to this sms to claim your prize.	0	2	0	0	1	3	0	1	0	1	0	0	0	0	0	0	0	0	0	0
Don’t worry, I guess he is busy.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Going for dinner. Msg you later.	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
Ok, I will call u up when I get some cash.	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Nistha Adhikari

	2	is	yours	this	claim	your	prize	I	am	pleased	tell	that	are	awarded	with	1500	bonus	Don't	worry
Congrats,You have won!!reply to our sms for a free nokia mobile + free camcorder.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Congrats! 1 year special cinema pass for 2 is yours. reply to this sms to claim your prize.	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
I am pleased to tell you that you are awarded with a 1500 Bonus Prize, reply to this sms to claim your prize.	0	0	0	1	1	1	2	1	1	1	1	1	1	1	1	1	1	0	0
Don't worry, I guess he is busy.	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	1
Going for dinner. Msg you later.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ok, I will call u up when I get some cash.	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0

Nistha Adhikari

	he	guess	busy	going	dinner	msg	later	ok	will	call	up	when	get	some	cash	label
Congrats,You have won!!reply to our sms for a free nokia mobile + free camcorder.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	+
Congrats! 1 year special cinema pass for 2 is yours. reply to this sms to claim your prize.	0	0	1	1	1	0	1	1	0	1	0	0	0	0	0	+
I am pleased to tell you that you are awarded with a 1500 Bonus Prize, reply to this sms to claim your prize.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	+
Don't worry, I guess he is busy.	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	-
Going for dinner. Msg you later.	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	-
Ok, I will call you up when I get some cash.	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	-

Nistha Adhikari

Positive Documents

	Congrats	you	have	won	repl y	to	ou r	sms	for	a	free	nokia	mobile	+	camcorder	1	year	special	cinema	pass
Congrats, you have won!!reply to our sms for a free nokia mobile + free camcorder.	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1	0	0	0	0	0
Congrats! 1 year special cinema pass for 2 is yours. reply to this sms to claim your prize.	1	0	0	0	1	2	0	1	1	0	0	0	0	0	0	1	1	1	1	1
I am pleased to tell you that you are awarded with a 1500 Bonus Prize, reply to this sms to claim your prize.	0	2	0	0	1	3	0	1	0	1	0	0	0	0	0	0	0	0	0	0

	2	is	yours	this	claim	your	prize	I	am	pleased	tell	that	are	awarded	with	1500	bonus	Don't	worry
Congrats,You have won!!reply to our sms for a free nokia mobile + free camcorder.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Congrats! 1 year special cinema pass for 2 is yours. reply to this sms to claim your prize.	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
I am pleased to tell you that you are awarded with a 1500 Bonus Prize, reply to this sms to claim your prize.	0	0	0	1	1	1	2	1	1	1	1	1	1	1	1	1	1	0	0

	he	guess	busy	going	dinner	msg	later	ok	will	call	up	when	get	some	cash	label
Congrats,You have won!!reply to our sms for a free nokia mobile + free camcorder.	0	0	0	0	0	0	0	0		0	0	0	0	0	0	+
Congrats! 1 year special cinema pass for 2 is yours. reply to this sms to claim your prize.	0	0	1	1	1	0	1	1	0	1	0	0	0	0	0	+
I am pleased to tell you that you are awarded with a 1500 Bonus Prize, reply to this sms to claim your prize.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	+

p (+) = 3/6

Using,
$p\ (wk|+) = nk+1/(n+|Vocabulary|)$,

Where, |vocabulary| = 54, n = 63, *nk* is the number of times the word,

p(**Congrats**|+) = 0.02564102564102564	p(**prize**|+) 0.017094017094017096
p(**you**|+) = 0.03418803418803419	p(**I**|+) 0.017094017094017096
p(**have**|+) = 0.017094017094017096	p(**am**|+) 0.02564102564102564
p(**won**|+) = 0.017094017094017096	p(**pleased**|+) 0.02564102564102564
p(**reply**|+) = 0.03418803418803419	p(**tell**|+) 0.02564102564102564
p(**to**|+) = 0.05982905982905983	p(**that**|+) 0.03418803418803419
p(**our**|+) = 0.017094017094017096	p(**are**|+) 0.017094017094017096
p(**sms**|+) = 0.03418803418803419	p(**awarded**|+) 0.017094017094017096

p(**for**|+) = 0.02564102564102564	p(**with**|+) 0.017094017094017096
p(**a**|+) = 0.02564102564102564	p (**1500**|+) 0.017094017094017096
p(**free**|+) = 0.02564102564102564	p(**bonus**|+) 0.017094017094017096
p(**nokia**|+) = 0.017094017094017096	p(**don't**|+) 0.017094017094017096
p(**mobile**|+) = 0.017094017094017096	p(**worry**|+) 0.017094017094017096
p (+|+) = 0.017094017094017096	p(**guess**|+) 0.017094017094017096
p(**camcorder**|+) = 0.017094017094017096	p(**he**|+) 0.017094017094017096
p (**1**|+) = 0.017094017094017096	p(**busy**|+) 0.017094017094017096
p(**year**|+) = 0.017094017094017096	p(**going**|+) 0.008547008547008548
p(**special**|+) = 0.017094017094017096	p(**dinner**|+) 0.008547008547008548
p(**cinema**|+) = 0.017094017094017096	p(**msg**|+) 0.008547008547008548
p(**pass**|+) = 0.017094017094017096	p(**later**|+) 0.008547008547008548
p(**for**|+) 0.017094017094017096	p(**ok**|+) 0.017094017094017096
p (**2**|+) 0.008547008547008548	p(**call**|+) 0.017094017094017096
p(**is**|+) 0.008547008547008548	p(**up**|+) 0.017094017094017096
p(**yours**|+) 0.008547008547008548	p(**when**|+) 0.008547008547008548
p(**this**|+) 0.008547008547008548	p(**get**|+) 0.017094017094017096
p(**claim**|+) 0.008547008547008548	p(**some**|+) 0.017094017094017096
p(**your**|+) 0.017094017094017096	p(**cash**|+) 0.008547008547008548

Nistha Adhikari

Negative documents

	Congrats	you	have	won	repl y	to	ou r	sms	for	a	free	nokia	mobile	+	camcorder	1	year	special	cinema	pass
Don't worry, I guess he is busy.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Going for dinner. Msg you later.	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
Ok, I will call u up when I get some cash.	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

	2	is	yours	this	claim	your	prize	I	am	pleased	tell	that	are	awarded	with	1500	bonus	Don't	worry
Don't worry, I guess he is busy.	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	1
Going for dinner. Msg you later.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ok, I will call u up when I get some cash.	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0

	he	guess	busy	going	dinner	msg	later	ok	will	call	up	when	get	some	cash	label
Don't worry, I guess he is busy.	1	1	1	0	0	0	0	0		0	0	0	0	0	0	-
Going for dinner. Msg you later.	0	0	0	1	1	1	1	0		0	0	0	0	0	0	-
Ok, I will call you up when I get some cash.	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	-

p (-) = 3/6

Using,
$p\ (wk|+) = nk+1/(n+|Vocabulary|)$,

Where, |vocabulary| = 54, n = 24, *nk* is the number of times the word,

p(**Congrats**|-) = 0. 01282051282051282	p(**prize**|-) 0. 01282051282051282
p(**you**|-) = 0. 038461538461538464	p(**I**|-) 0. 05128205128205128
p(**have**|-) = 0. 01282051282051282	p(**am**|-) 0. 01282051282051282
p(**won**|-) = 0. 01282051282051282	p(**pleased**|-) 0. 01282051282051282
p(**reply**|-) = 0. 01282051282051282	p(**tell**|-) 0. 01282051282051282
p(**to**|-) = 0. 01282051282051282	p(**that**|-) 0. 01282051282051282
p(**our**|-) = 0. 01282051282051282	p(**are**|-) 0. 01282051282051282
p(**sms**|-) = 0. 01282051282051282	p(**awarded**|-) 0. 01282051282051282
p(**for**|-) = 0. 01282051282051282	p(**with**|-) 0. 01282051282051282
p(**a**|-) = 0. 02564102564102564	p (**1500**|-) 0. 01282051282051282
p(**free**|-) = 0. 01282051282051282	p(**bonus**|-) 0. 01282051282051282
p(**nokia**|-) = 0. 01282051282051282	p(**don't**|-) 0. 02564102564102564
p(**mobile**|-) = 0. 01282051282051282	p(**worry**|-) 0. 02564102564102564
p (+|-) = 0. 01282051282051282	p(**guess**|-) 0. 02564102564102564
p(**camcorder**|-) = 0. 01282051282051282	p(**he**|-) 0. 02564102564102564
p (**1**|-) = 0. 01282051282051282	p(**busy**|-) 0. 02564102564102564
p(**year**|-) = 0. 01282051282051282	p(**going**|-) 0. 02564102564102564
p(**special**|-) = 0. 01282051282051282	p(**dinner**|-) 0. 02564102564102564

p(**cinema**|-) = 0. 01282051282051282	p(**msg**|-) 0. 02564102564102564
p(**pass**|-) = 0. 01282051282051282	p(**later**|-) 0. 02564102564102564
p(**for**|-) 0. 01282051282051282	p(**ok**|-) 0. 02564102564102564
p (**2**|-) 0. 01282051282051282	p(**call**|-) 0. 02564102564102564
p(**is**|-) 0. 02564102564102564	p(**up**|-) 0. 02564102564102564
p(**yours**|-) 0. 01282051282051282	p(**when**|-) 0. 02564102564102564
p(**this**|-) 0. 01282051282051282	p(**get**|-) 0. 02564102564102564
p(**claim**|-) 0. 01282051282051282	p(**some**|-) 0. 02564102564102564
p(**your**|-) 0. 01282051282051282	p(**cash**|-) 0. 02564102564102564

Classifying the first sentence as + or -:

"I am busy. I will msg you later."

$y+ = p\ (+)\ p(I|+)\ p(\text{am}|+)\ p(\text{busy}|+)\ p(\text{I}|+)\ p(\text{will}|+)\ p(\text{msg}|+)\ p(\text{you}|+)\ p(\text{later}|+)$ = 1.0662252081783243e-13
$y-= p\ (+)\ p(I|+)\ p(\text{am}|+)\ p(\text{busy}|+)\ p(\text{I}|+)\ p(\text{will}|+)\ p(\text{msg}|+)\ p(\text{you}|+)\ p(\text{later}|+)$ = 5.465237180357796e-12

The sentence will be classified as not spam.

Classifying the second sentence as + or -:

"Congrats! You are awarded a free mobile."

$y+ = p\ (+)\ p\ (\text{Congrats}\ |+)\ p\ (\text{You}\ |+)\ p\ (\text{are}\ |+)\ p\ (\text{awarded}\ |+)\ p\ (\text{a}\ |+)\ p\ (\text{free}\ |+)\ p\ (\text{mobile}\ |+)$ = 1.4394040310407373e-12
$y-= p\ (+)\ p\ (\text{Congrats}\ |+)\ p\ (\text{You}\ |+)\ p\ (\text{are}\ |+)\ p\ (\text{awarded}\ |+)\ p\ (\text{a}\ |+)\ p\ (\text{free}\ |+)\ p\ (\text{mobile}\ |+)$ = 8.539433094309056e-14

The sentence will be classified as spam.