



Word2Vec:

→ one of the popular method of word embedding representation with contextual information.

→ 2 types of Architecture

↓
Skipgram

↓
CBOW

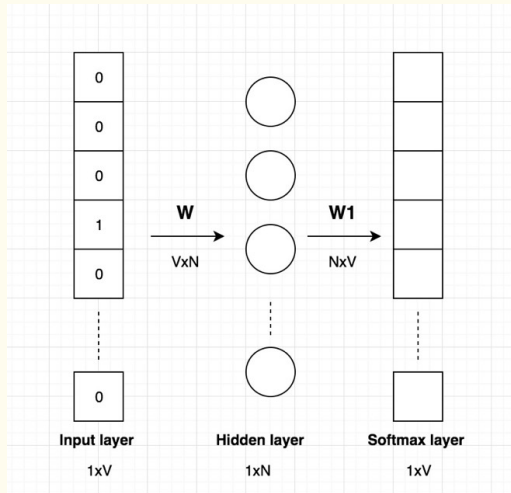
↓
given an input word in a sentence, the network will predict how likely it is for each word in the vocabulary being that input word's nearby word.

↓
given a context of words (surrounding a word) in a sentence, the Network will predict how likely it is for each word in the vocabulary being that word

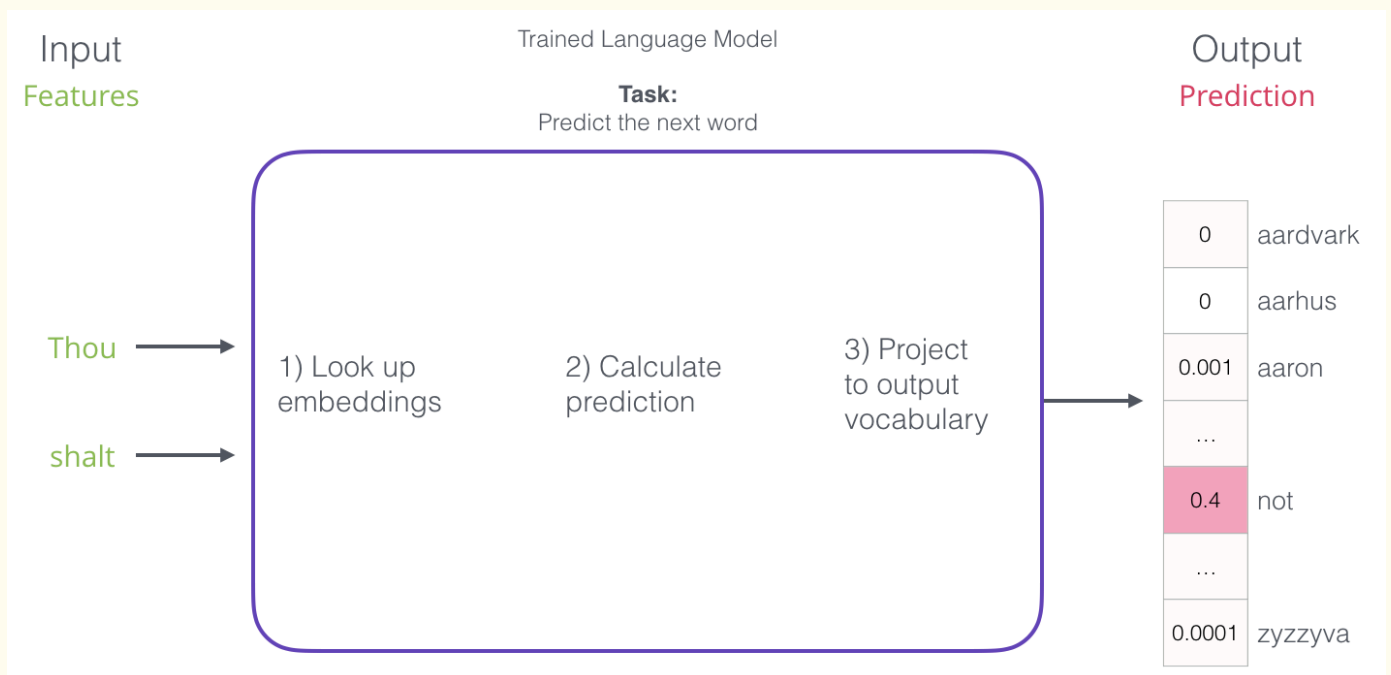
Objective: similar words should be as close as possible and opposite word should be as distant as possible in embedding dimensional space.



skip gram:



Sentence	Training examples
He says make America great again.	(he,says), (he,make)
He says make America great again.	(says,he), (says,make), (says,america)
He says make America great again.	(make,he), (make,says), (make,america), (make,great)
He says make America great again.	(america,says),(america,make) (america,great),(america,again)
He says make America great again.	(great,make), (great,america),(great,again)
He says make America great again .	(again,america), (again,great)



Thou shalt not make a machine in the likeness of a human mind

Sliding window across running text

input 1	input 2	output
thou	shalt	not
shalt	not	make
not	make	a
make	a	machine
a	machine	in

Thou shalt not make a machine in the likeness of a human mind

input word	target word
not	thou
not	shalt
not	make
not	a
not	machine
not	in
not	the
not	...
make	not
make	shalt
make	make
make	a
make	machine
make	in
make	the
make	...
a	not
a	shalt
a	make
a	a
a	machine
a	in
a	the
a	...
machine	not
machine	shalt
machine	make
machine	a
machine	machine
machine	in
machine	the
machine	...
in	not
in	shalt
in	make
in	a
in	machine
in	in
in	the
in	...

instead of looking 2 words before target, look 2 words after it also.

WS=2

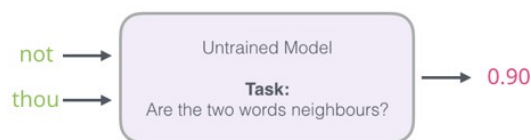


Negative Sampling:

Change Task from



To:



input word	target word
not	thou
not	shalt
not	make
not	a
make	shalt
make	not
make	a
make	machine

input word	output word	target
not	thou	1
not	shalt	1
not	make	1
not	a	1
make	shalt	1
make	not	1
make	a	1
make	machine	1

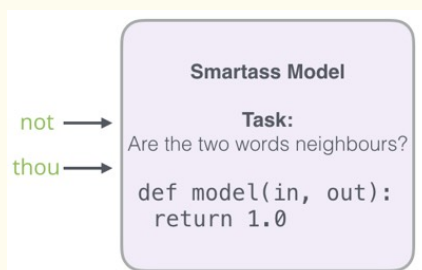
→ make new target var

→ if neighbour = 1
not neighbour = 0

problem:

→ if our all target = 1

↳ model accuracy = 100%
but model won't be learning anything.



→ so we need to impute negative samples.

→ randomly sample words from vocab for negative sampling.

input word	output word	target
not	thou	1
not	aaron	0
not	taco	0
not	shalt	1
not	make	1

Pick randomly from vocabulary (random sampling)

Word	Count	Probability
aardvark		
aarhus		
aaron		
taco		
thou		
zyzzyva		

Training:

Skipgram

shalt	not	make	a	machine
input		output		
make		shalt		
make		not		
make		a		
make		machine		

Negative Sampling

input word	output word	target
make	shalt	1
make	aaron	0
make	taco	0

Embedding

	aardvark
	aarhus
	aaron
	...
	not
	...
	zyzzyva

Context

	aardvark
	aarhus
	aaron
	...
	taco
	...
	thou
	...
	zyzzyva

Look up embeddings

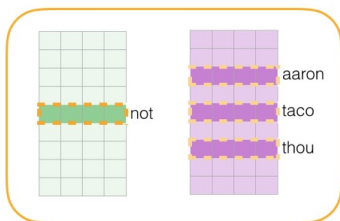


aaron
taco
thou

input word	output word	target	input • output
not	thou	1	0.2
not	aaron	0	-1.11
not	taco	0	0.74

input word	output word	target	input • output	sigmoid()
not	thou	1	0.2	0.55
not	aaron	0	-1.11	0.25
not	taco	0	0.74	0.68

input word	output word	target	input • output	sigmoid()	Error
not	thou	1	0.2	0.55	0.45
not	aaron	0	-1.11	0.25	-0.25
not	taco	0	0.74	0.68	-0.68



Update Model Parameters

Actual Target

0
0
0
...
0
1
...
0

not



Update Model Parameters

Model Prediction

0	aardvark	0
0	aarhus	0
0.001	aaron	-0.001
...
0.4	taco	-0.4
0.001	thou	0.999
...
0.0001	zyzzyva	-0.0001

Error

Hyperparameter:

→ window size

→ # of negative samples

Negative samples: 2

input word	output word	target
make	shalt	1
make	aaron	0
make	taco	0

Negative samples: 5

input word	output word	target
make	shalt	1
make	aaron	0
make	taco	0
make	linglonger	0
make	plumbus	0
make	mango	0

→ In gluSiu w2v default:

WS = 5

NS = 5



CBOW:

Jay was hit by a _____ bus in...

by	a	red	bus	in
----	---	-----	-----	----

input 1	input 2	input 3	input 4	output
by	a	bus	in	red