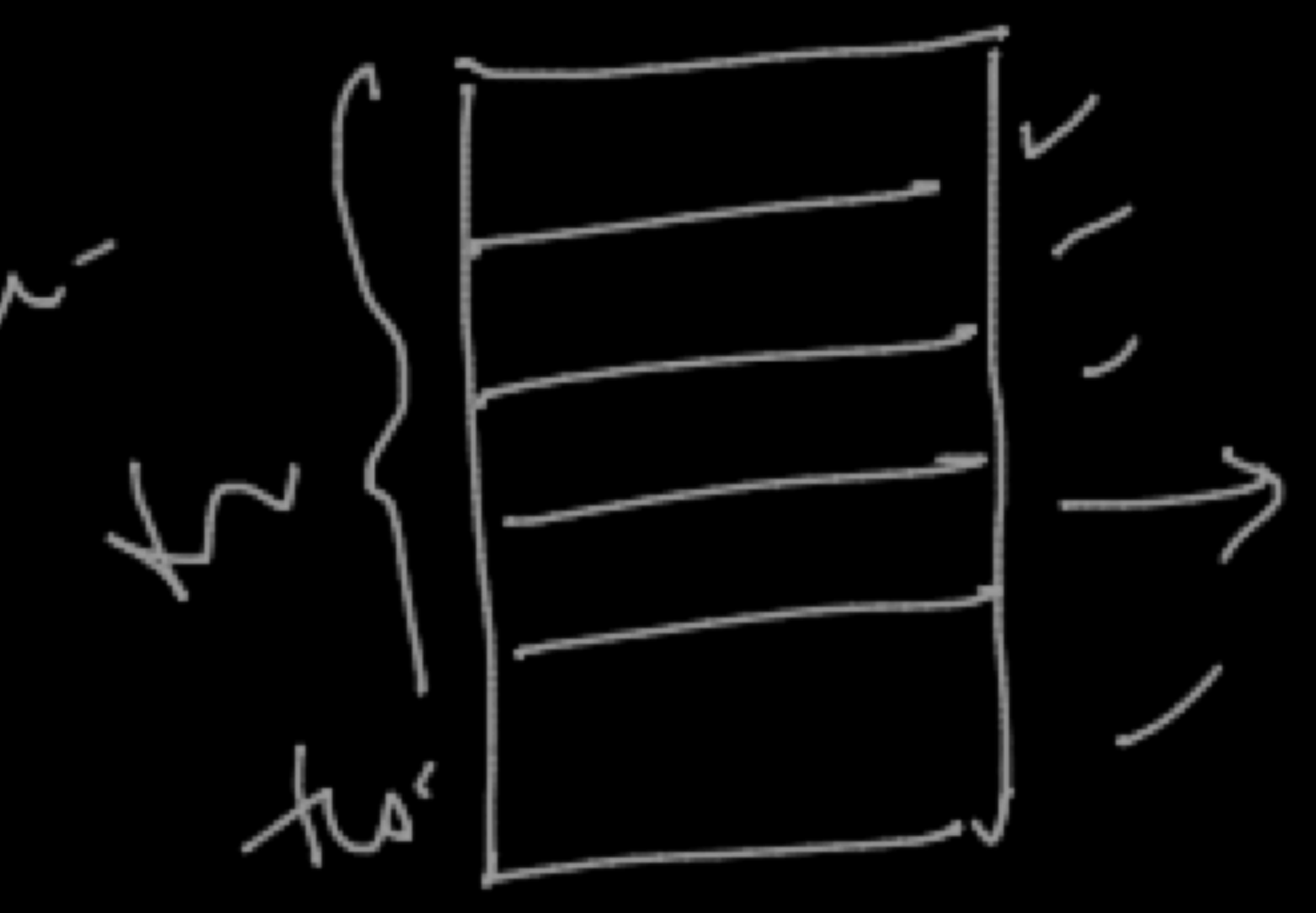
K = 1, 2, 3

V = 10, 20, 30 .

→ Sklearn



model (K = 1, V = 10)

model.fit(train)

model.predict(testing) → training Error testing Error

sklearn

keras

keras.

kc = kerasclassifier (keras)

kc.fit ( )

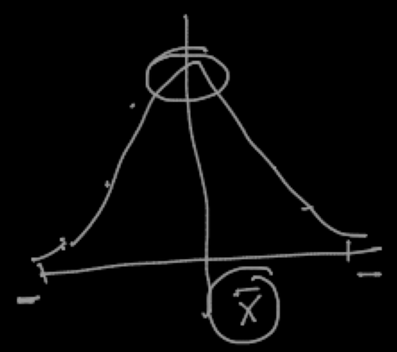
kc.predict ( ).

Gridsearchcvc)

	W_1	W_2	W_3	W_4	
→	10	5	2.3	11	→ .
	-	-	.	.	.
	-	-	.	-	.
	0	0	0	0	

$\hat{y}$

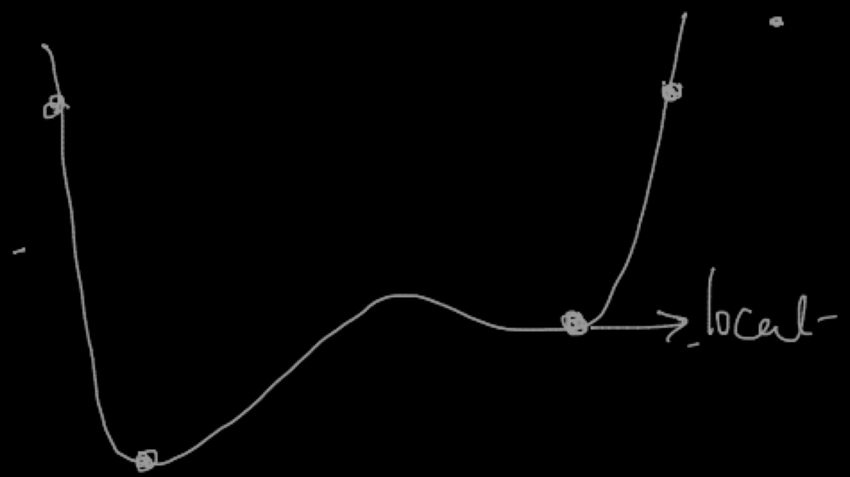
np.random.randn(10)



1. Gradient Descent -
   - Forward prop. (activation fn)
   - Back prop. (Chain Rule - weight update).
   - Epochs (Batch Size).
   - learning rate.

optimizers. ←
- Momentum
- Nesterov momentum
- Adam
- RMS prop.

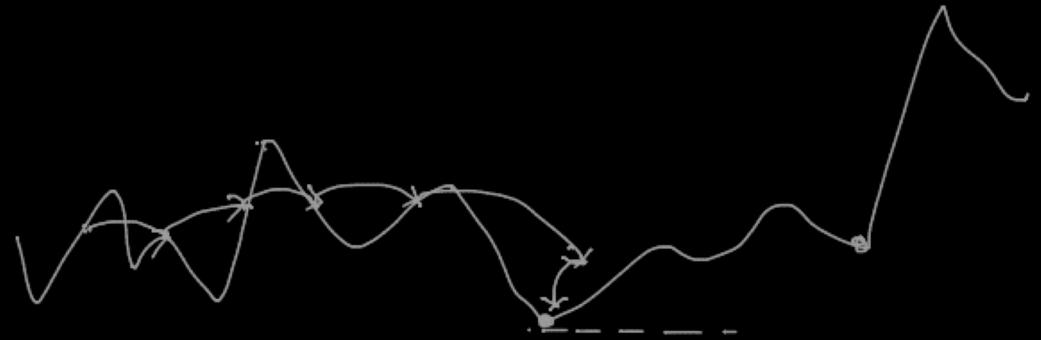
1. local minima
2. Exploding / Vanishing gradients.
3. Saddle point -



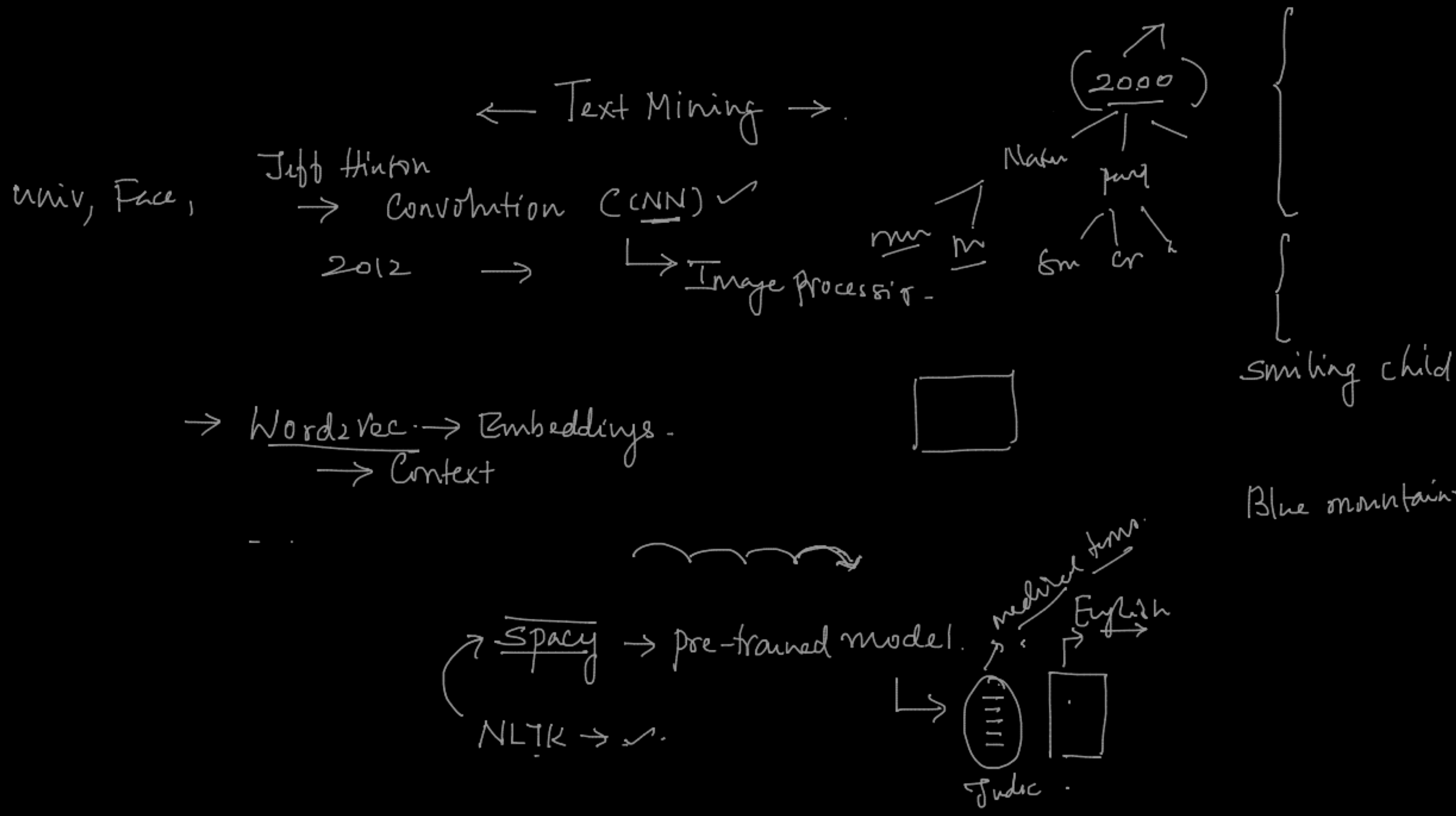
$$\Rightarrow \left(\frac{\partial L}{\partial \cdot}\right)\left(\frac{\partial}{\partial}\right)\left(\frac{\partial}{\partial}\right)(\cdot)\cdots$$

Weight → 0

Δw.



PowerPoint File Edit View Insert Format Arrange Tools Slide Show Window Help



First step → Convert text to vectors
└ Classification

R1 → x_1 x_2 x_3 ... | y
(- - - - -) | P
- - - - - | N.
- - - - - | -

1. Bag of Words } 'n-grams'
2. Binary BoW }
3. TF-IDF

— Term Freq — Inverse Doc Freq.

For $n=1$; uni-gram
$n=2$; bi-gram
$n=3$; tri-gram.

1. Bag of Words

Corpus

D1: The car is driven on the road.

D2: The truck is driven on the highway

Vocab: The car truck is driven on road highway

R1 :

2	1	0	1	1	1	1	0

R2 :

2	0	1	1	1	1	0	1

Regular Exp → [a-z][A-Z][0-9] →

PowerPoint File Edit View Insert Format Arrange Tools Slide Show Window Help

7 / 9

98%.
78%. → Good accuracy

Corpus { D1: Tom is (not clever) but he is honest
D2: Jon is Clever but not honest.

Vocab. Tom | Jon | is clever honest not but he → unigram.

n → 1, 2, 3, ...

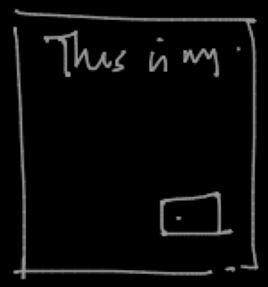
unigram → 1 w
n-gram.
↳ bi-gram. (2)
tri-gram.

not honest	but not	Jon is	Tom is	is not	not clever	clever but	but he	he is	is honest	is clever
D1: →	0	0	0	1	1	1	1	1	1	0
D2: →	1	1	1	0	0	0	1	0	0	0	0

	x_1	x_2	x_3	x_4	x_5	...
D1	0	0	0	1	1	1	...
D2	1	1				

GPT3 → 1 Trillion →

Pin



→ 3000 Reviews.