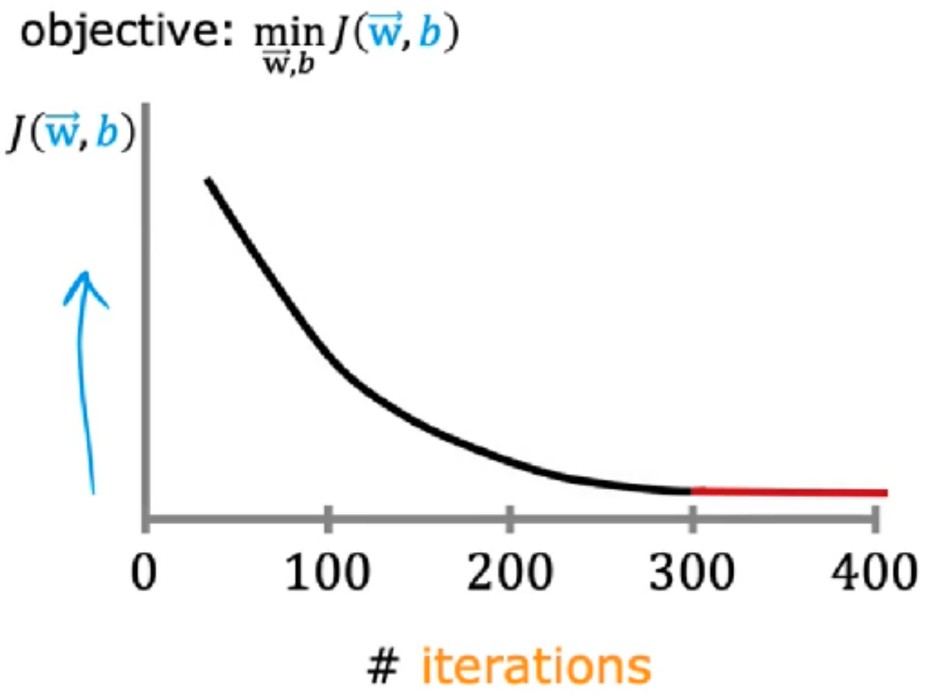
3_Checking Gradient Descent for Convergence

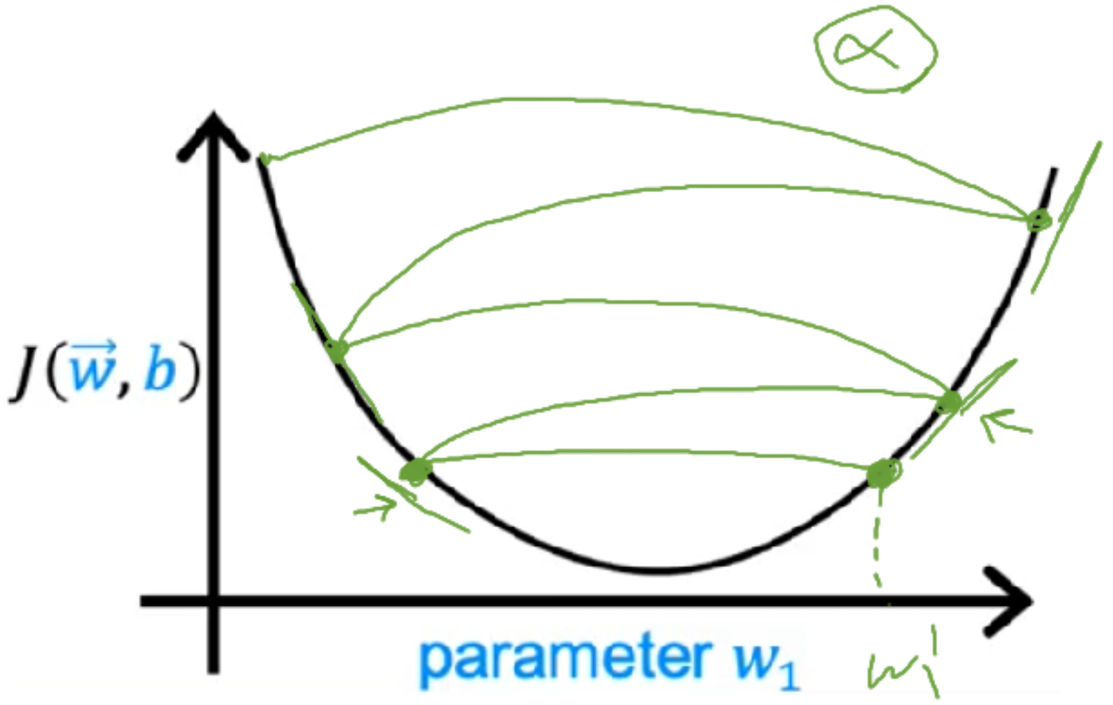
Make sure gradient descent is working correctly

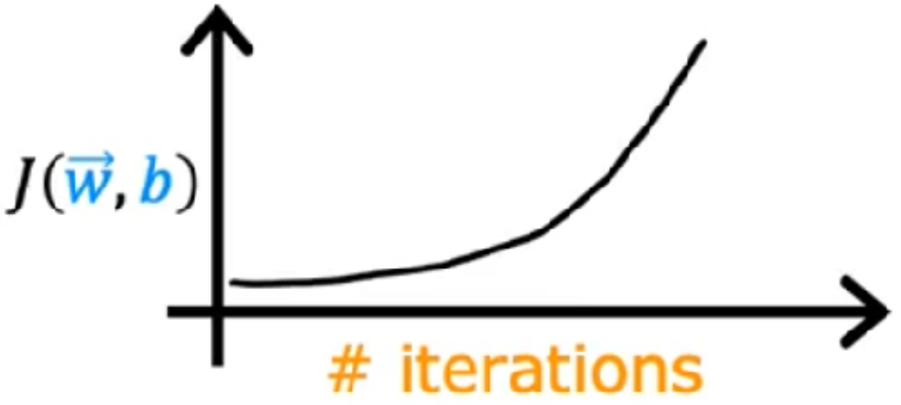


4. Choosing the Learning Rate

$w^{new} = w^{old} - \alpha \frac{\partial}{\partial w} J(w)$

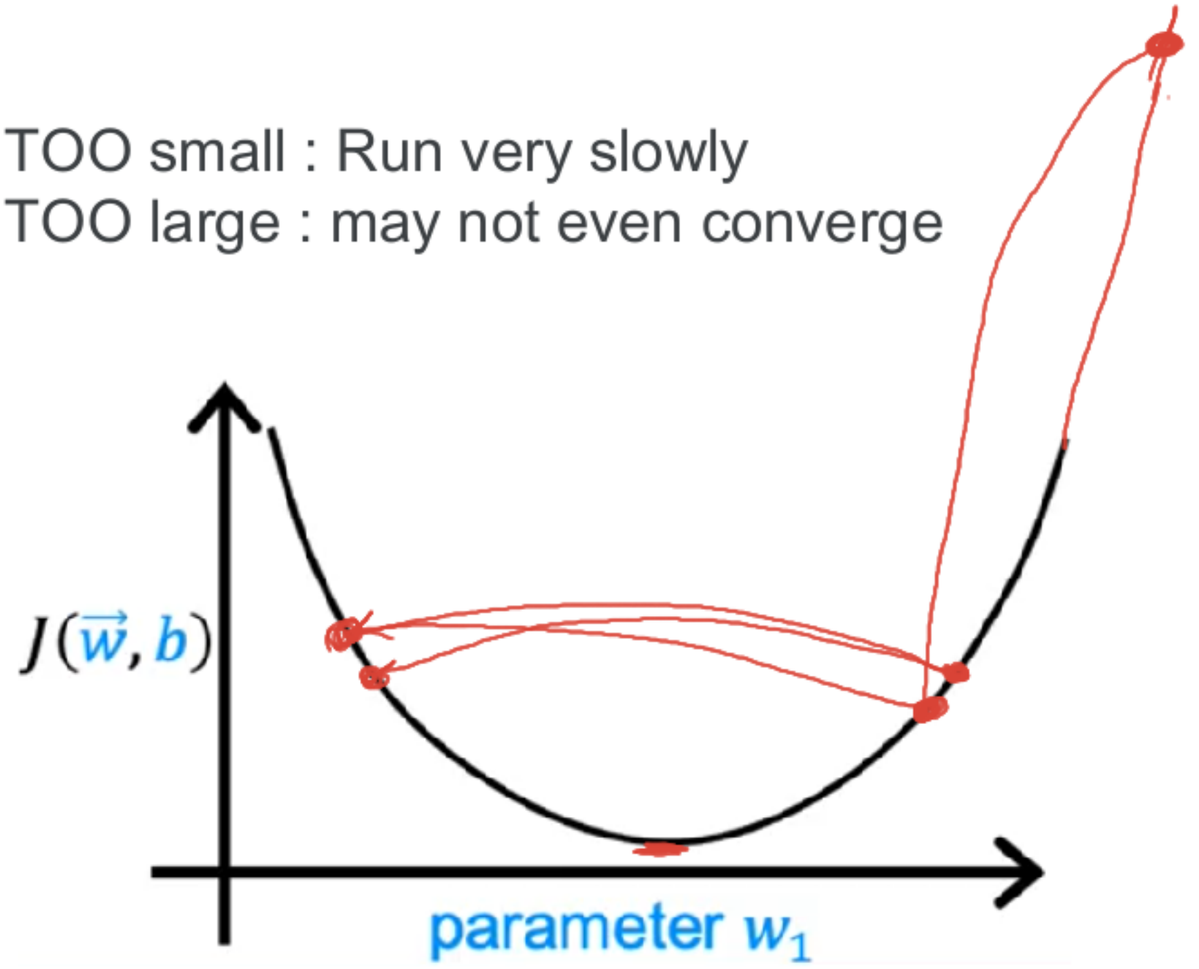
TOO small : Run very slowly
TOO large : may not even converge

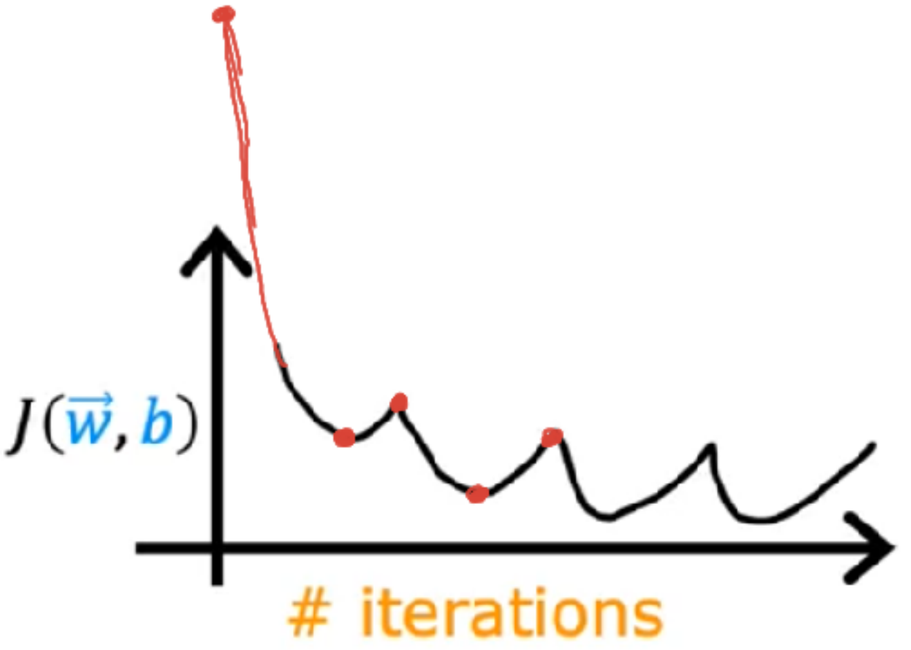




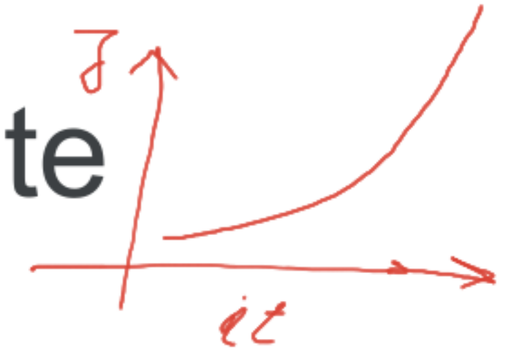
4. Choosing the Learning Rate

TOO small : Run very slowly
TOO large : may not even converge



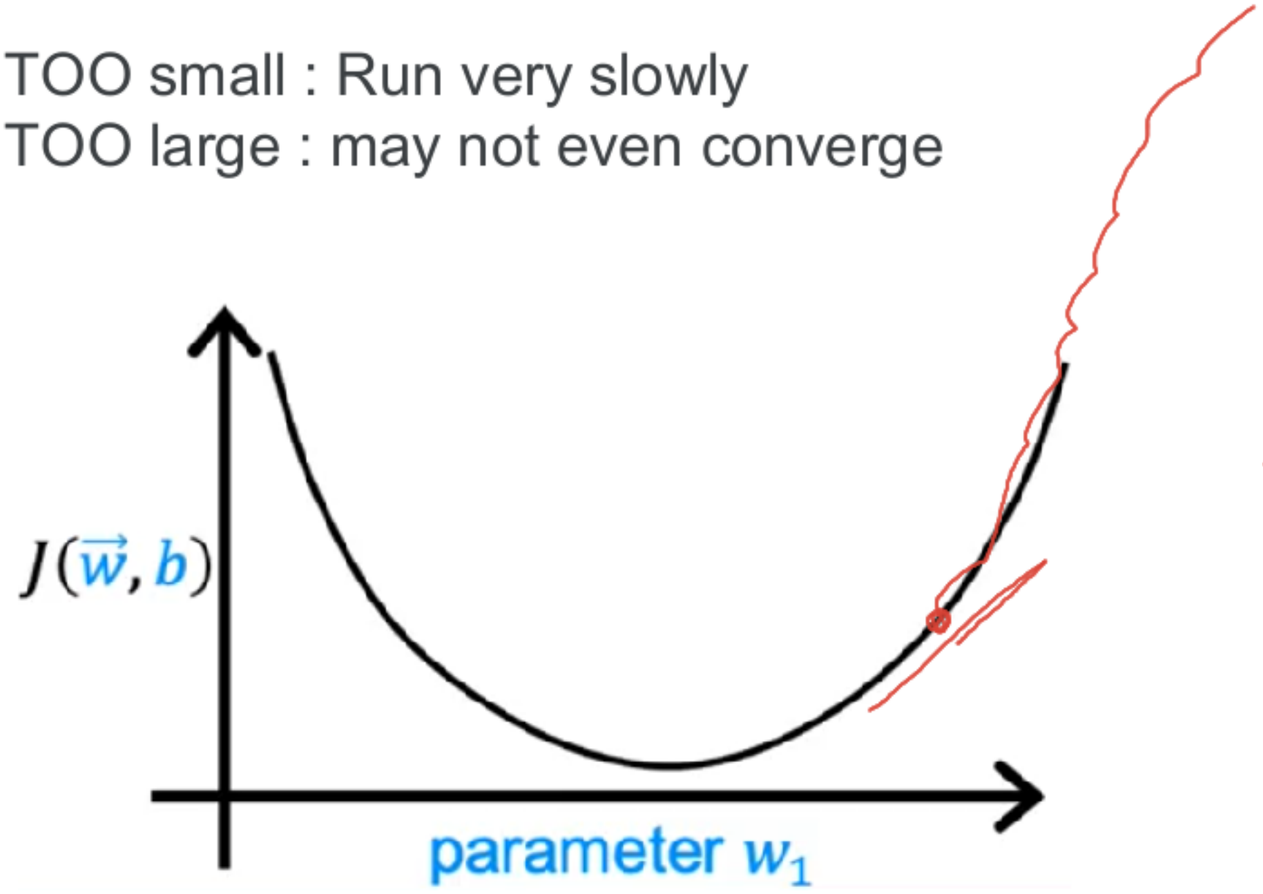


4. Choosing the Learning Rate



TOO small : Run very slowly
TOO large : may not even converge

$$w^{new} = w^{old} - \alpha \frac{\partial J}{\partial w}$$



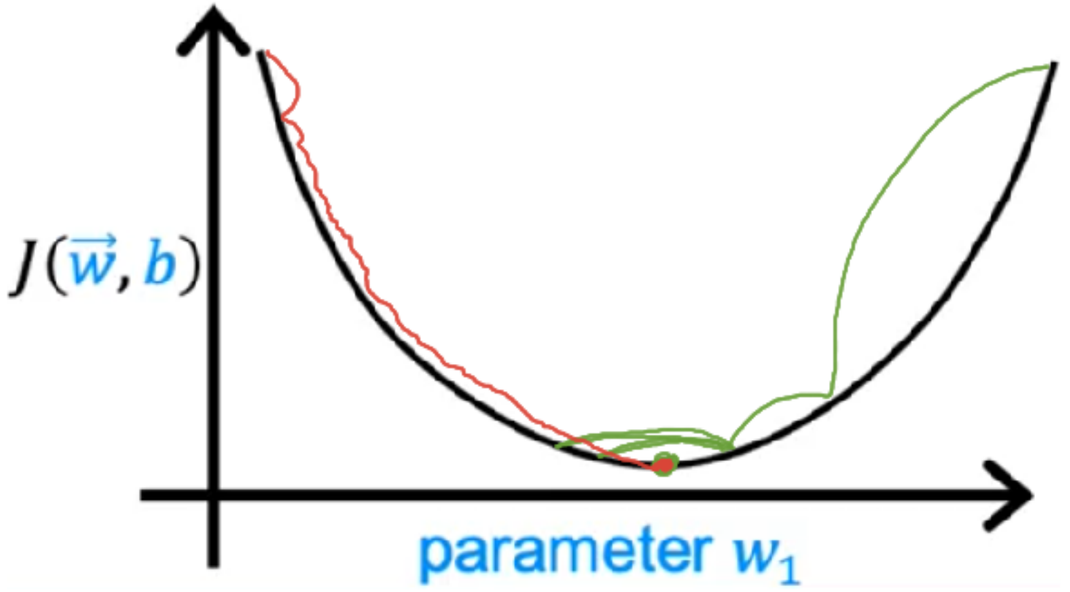
$$w_1 = w_1 + \alpha d_1$$ ☹

use a minus sign

$$w_1 = w_1 - \alpha d_1$$ ☺

4. Choosing the Learning Rate

TOO small : Run very slowly
TOO large : may not even converge

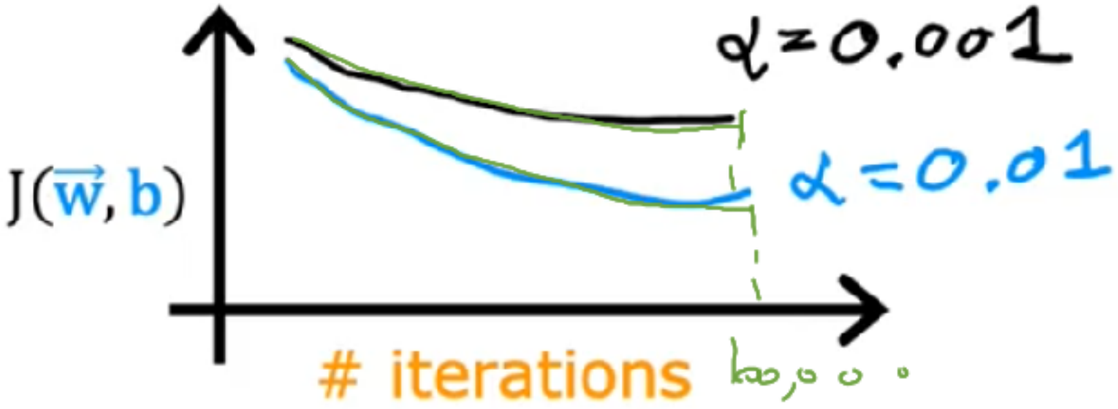


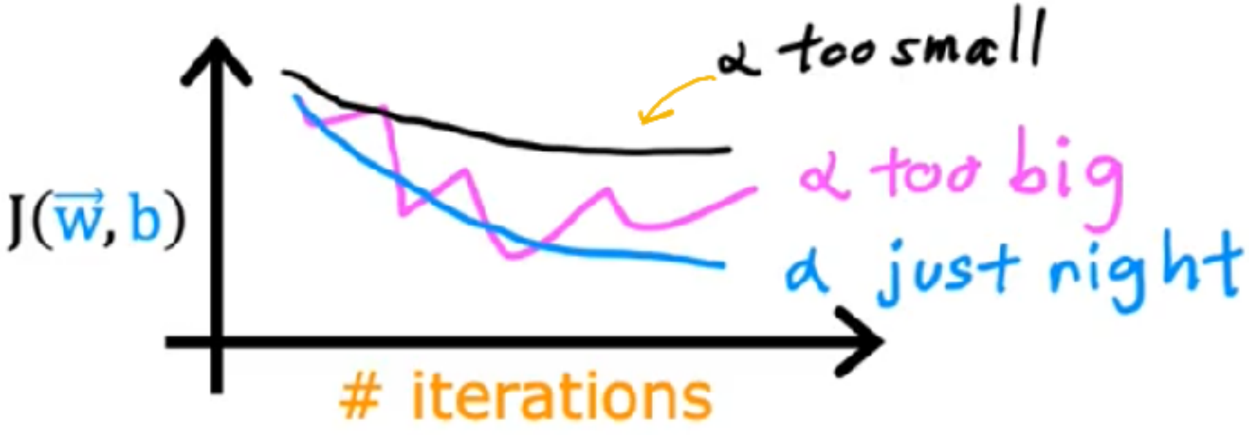
If α is too small, gradient descent takes a lot more iterations to converge

Values of α to try:

... 0.001 0.003 0.01 0.03 0.1 0.3 1 ...

3X ≈3X 3X ≈3X 3X ≈3X



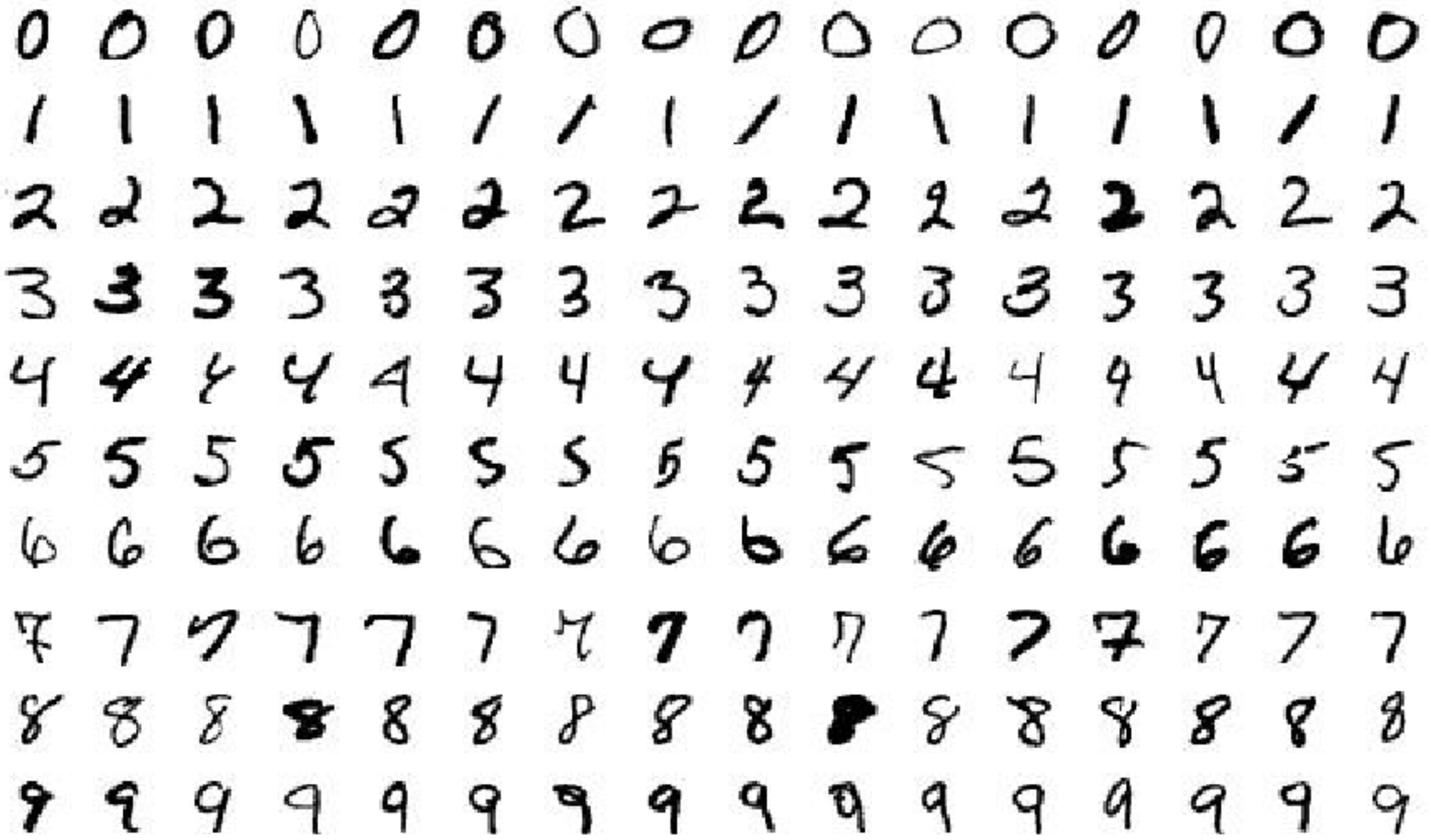


Feature Engineering

Choosing the right features is a critical step to making the algorithm work well.

MNIST Dataset

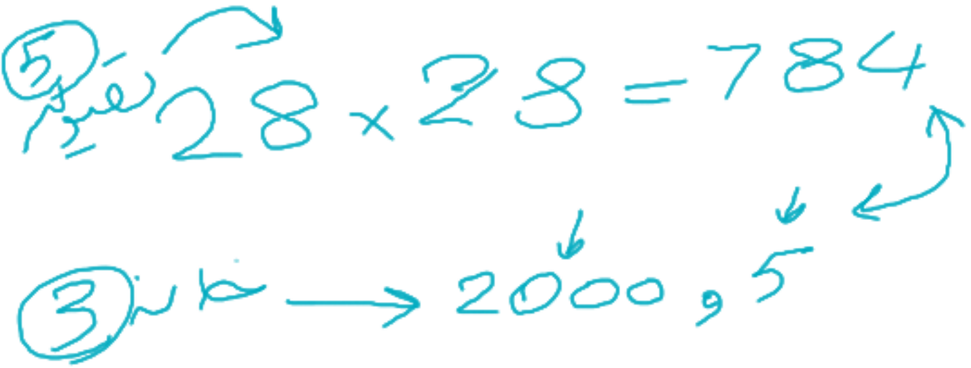
60,000 training images
10,000 testing images.



MNIST Dataset

60,000 training images
10,000 testing images.

?



	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	29	150	195	254	255	254	176	193	150	96	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	48	166	224	253	253	234	196	253	253	253	253	233	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	93	244	249	253	187	46	10	8	4	10	194	253	253	233	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	107	253	253	230	48	0	0	0	0	0	192	253	253	156	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	3	20	20	15	0	0	0	0	0	43	224	253	245	74	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	249	253	245	126	0	0	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	0	14	101	223	253	248	124	0	0	0	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0	11	166	239	253	253	253	187	30	0	0	0	0	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0	0	16	248	250	253	253	253	253	232	213	111	2	0	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	0	0	0	0	43	98	98	208	253	253	253	253	187	22	0	0	0	0	0	0	0
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	51	119	253	253	253	76	0	0	0	0	0	0	0
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	183	253	253	139	0	0	0	0	0	0	0
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	182	253	253	104	0	0	0	0	0	0	0
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	85	249	253	253	36	0	0	0	0	0	0	0
19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	60	214	253	253	173	11	0	0	0	0	0	0	0
20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	98	247	253	253	226	9	0	0	0	0	0	0	0	0
21	0	0	0	0	0	0	0	0	0	0	0	0	42	150	252	253	253	233	53	0	0	0	0	0	0	0	0	0
22	0	0	0	0	0	0	42	115	42	60	115	159	240	253	253	250	175	25	0	0	0	0	0	0	0	0	0	0
23	0	0	0	0	0	0	187	253	253	253	253	253	253	253	197	86	0	0	0	0	0	0	0	0	0	0	0	0
24	0	0	0	0	0	0	103	253	253	253	253	253	232	67	1	0	0	0	0	0	0	0	0	0	0	0	0	0
25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0