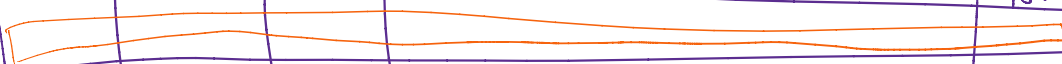


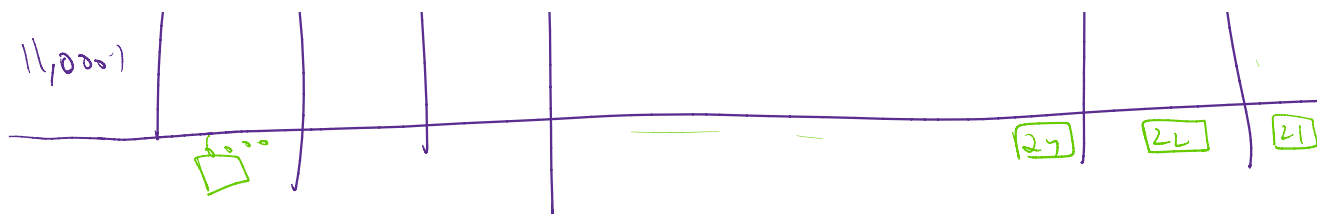
	Sentiment
1.) _____	1
2.) _____	0
3.) _____	1
⋮	⋮
14,677) _____	1

Randomly Pick
80 %
to be the
training data.

Review	Sentiment
1.) _____	1
2.) _____	0
3.) _____	1
⋮	⋮
11,000) _____	

← Training Data

	W_1	W_2	W_3	...	W_{1000}	W_{1001} W_{1002}
1.)						$(11,000 \times 1,000)$ ↓ a matrix of numbers
2.)						
3.)						
4.)						
⋮						
11,000)						



$$P(\text{sentiment} = 1) = \frac{1}{1 + e^{-(\alpha + \beta_1 w_1 + \beta_2 w_2 + \dots + \beta_{1000} w_{1000})}}$$

$$= \underline{0.7} \rightarrow 1$$

$$0.13 \rightarrow 0$$

TR₁ = [good phone. bad battery. - Narendra Gandhi]

good w_1 w_2 w_{1000}

→ [1 0 0 0 1 0 0 1 0 0 0 0 0 0 0] ←

(IDF score) [D D D D D D D D D D]

$$1 + \ln\left(\frac{1+n}{1+df}\right)$$

VADER

Input → Sentence(s)

Output: 'neg' + 'neu' + 'pos' = 1

compound → [-1, 1]

compound $\left\{ \begin{array}{l} [-1, -0.05) \rightarrow \text{'-ve' sentiment} \\ [-0.05, +0.05] \rightarrow \text{neutral sentiment} \\ (0.05, 1] \rightarrow \text{'+ve' sentiment} \end{array} \right.$