Word Embedding

→ DIFFERENT FROM CV & TF-IDF

1) My dog is brown in color
2) Cat is white
3) Dog and Cat are freinds
⋮
60,000 Rows.

↓

['my dog is brown in color',
'Cat is white',
'Dog and cat are Freinds',
⋮]

"corpus"

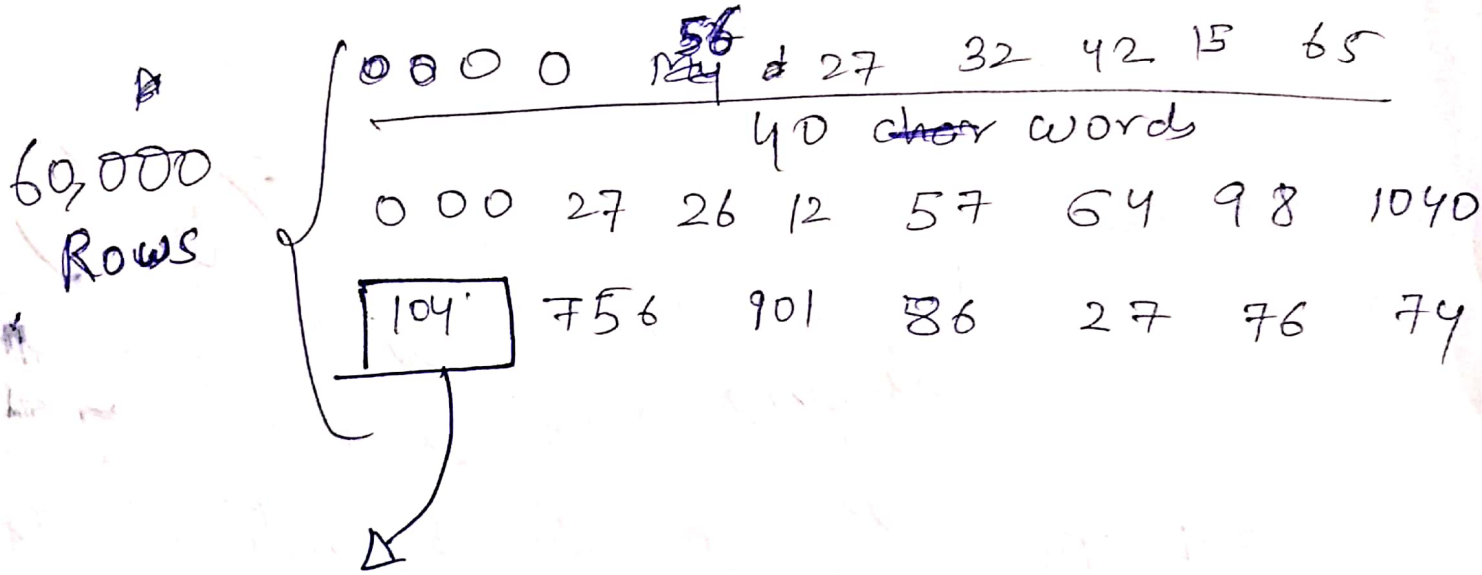
My dog is brown in color

0 0 0 0 0 9 56 27 32 42 15 65

"40 ~~char~~words" Fixed

↓

Padding



[_, _, _, _, _,
_, _, _, _, _, _,
_, _, _, _, _, _, _]

properties of a word.

LSTM Model

First layer

Embedding Layer (vocsize, embedding_vector_feature, input length = sent length)

10K → vocsize

100 → embedding_vector_feature

40 → input length = sent length

40, 100

LSTM Layer

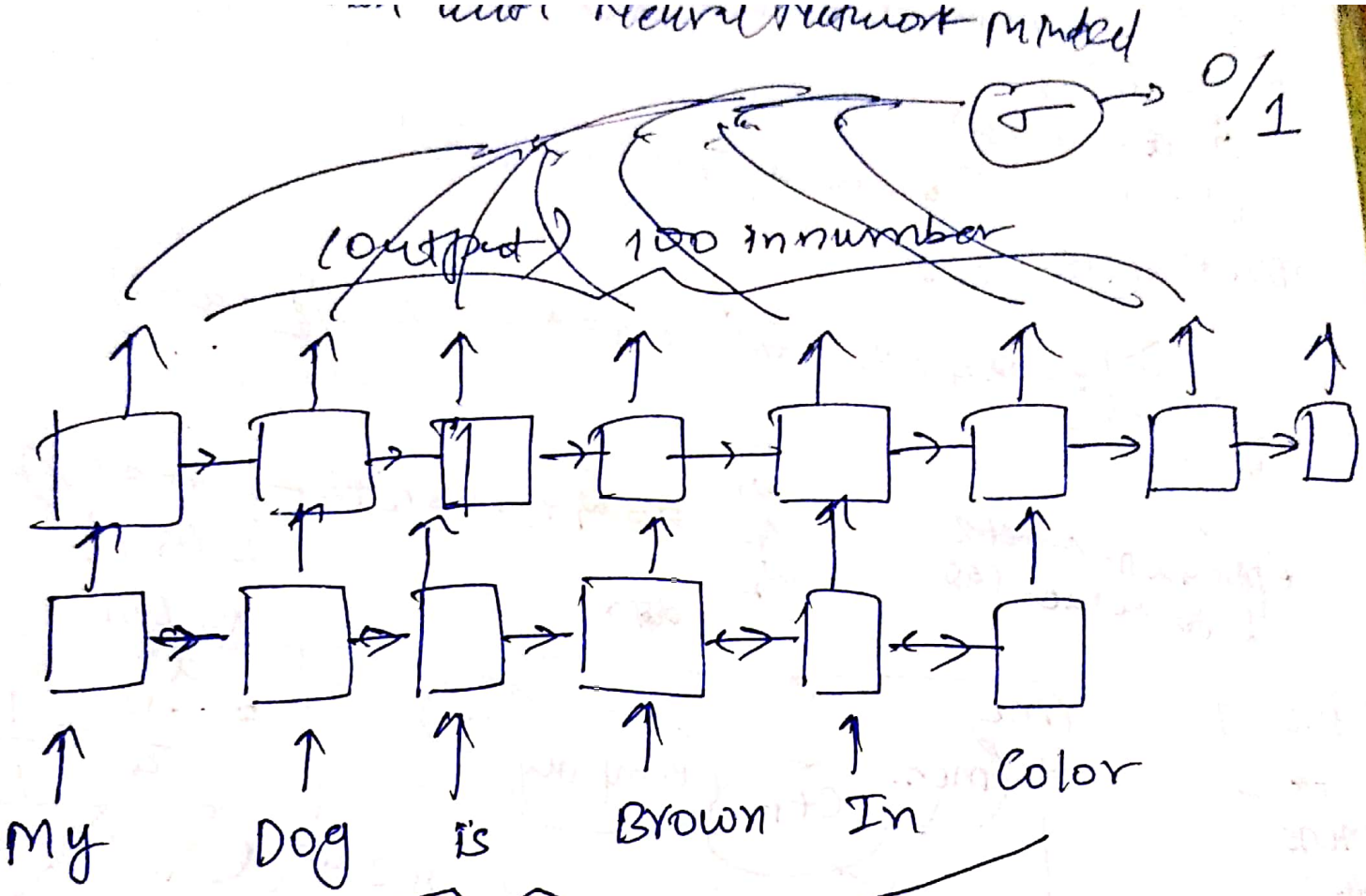
Training LSTM

~~Changing~~

Tuning Embedding Layer with Label

My dog is brown in color	0 Neg
Cat & white	0 Neg
Dog and Cat are Freinds	1 Pos

Trained Pattern will be implemented To to prediction



LSTM Working

$$Z = WX + B \rightarrow \text{Linearity}$$

$$A = \sigma(Z) \rightarrow \text{Non Linearity}$$