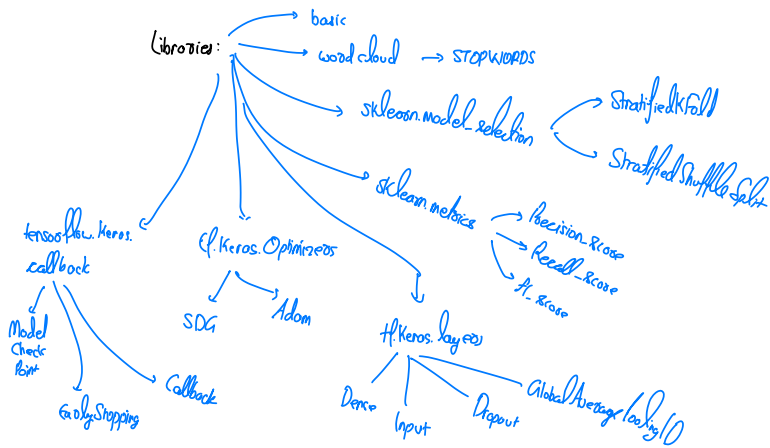


Goal: To find if a tweet was about disasters or not



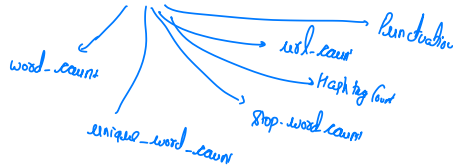
- Read (r): `pd.read_csv('file')`

- Missing data:
 - Keyword
 - location

- Step: Variative # of mixing entries of 'Location' & 'Keyword'
 - using \rightarrow $\text{isnull}()$
 - \rightarrow $\text{sum}()$

- Steps: find number of unique 'keyword' & 'location'

- Step 3: let add some extra features



- Plot a distribution plot

- for embeddings we are going to use GloVe to import using np load('')

- Steps: Expand the observation

Step 6: Its time for Model 5 we are using BERT

- Step 3: Encode using tokenizer, & train the model with it