- ***the order the words appear in***
  - bag of words ignored the order which the words appear in
  - each word in the block of text has an index
    - -> where in the block that word is
    - -> e.g there are 100 words in the block of text, this is the 5th word
    - -> this is less effective for huge vocabularies
      - -> when the order of the words change then the meaning of the text does
      - -> we don't have ways of grouping the words
- ***word embedding***
  - -> translating the words into a vector
  - -> the vector has n dimensions
  - -> the word "good" is represented by a vector in 3D cartesian space
    - -> if two words have different meanings then they will point in different directions in the space
    - -> this is what it is trying to do
      - vector representation of a word
    - -> the angle between those vectors
    - -> each of the word embeddings is a layer -> it can learn the word embeddings in each of the layers
  - -> the models are trained -> it's learnt the words
  - -> you can use pre-trained word embeddings
  - -> word embeddings are vectorised representations of words in a given document that places words with similar meanings near each other