

+ Embedding:

Words and structure of words are vectorized
to look for clusters.

So for movie reviews 2 distinct clusters
one for positive one for negative

+ Visualization (Σ M DB)

1. Load the data and extract the
tensor data.

2. Turning it to np arrays.

+ Have to use vector

Higher dimension clusters.

+ model \rightarrow Embedding \rightarrow Flatten \rightarrow Dense
 \downarrow
usually use a 1D pooling layer.

Global Average 1D Pooling \Rightarrow Average across
the vector \rightarrow flatten
it out.

Projector. Tensor flow. etc.

+ Returning to the Sarcasm data Set.

- We split the corpus.
- Tokenize
- Pad the tokenized sequences.
- Create MN

+ Loss function

more prediction, confidence fell, very common with text data.

+ Pre-Tokenized Datasets