**Word 2 Vec**

Word 2 Vec was introduced by Google in 2014. The paper was released in that year. It is used to convert the word into vectors. The same thing can be achieved with BOW and Tf-Idf however, the difference is that, Word 2 Vec gives the semantic meaning.

Suppose we have 3 words, Tasty, Delicious and Cricket. Now, word 2 vec converts those into 3 vectors and gives a Dense Matrix/Vector(non-zero) giving us the result that Tasty is similar to Delicious .

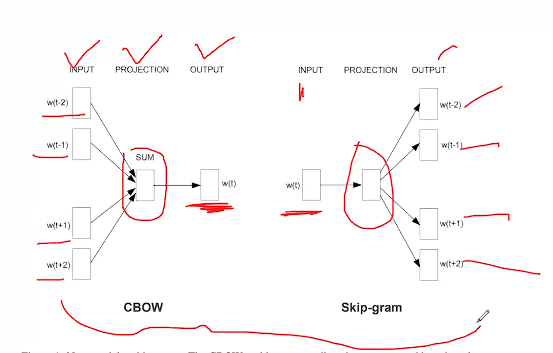
The calculation that takes place here is Matrix Factorization. It is a DL approach hence it is a Black-Box technique.

* If our W1(Tasty) and W2(Delicious) are semantically similar than their vectors V1 and V2 should also be closer.
* Relationship:- The relationship between Vman Vwoman is similar to Vking and Vqueen. Also, Vindia Vnewdelhi is similar to Vusa and Vwashingtondc.

Word 2 Vec is of two types:- CBOW(Continuous BOW) and SkipGram

CBOW – If I give my machine context words it is able to predict a Target Word/Focus Word. They do that with the help of MLP or DNN.

SkipGram – The Target here is to predict the Context word given the Focus word.



Word 2 Vec model is trained on Google News Corpus. Word 2 Vec uses Gensim Library.