# Skip-gram model

## ¿Qué es?

Es una arquitectura que, dado una secuencia de oraciones o corpus, se analiza las palabras en el en búsqueda del contexto de las palabras y la búsqueda de sus vecinos.

## ¿Como funciona?

Se utiliza una red neuronal simple con el objetivo de transformar las palabras a vectores (Word2Vect)

1. Se entrena con pares de palabras tomadas de algún documento

Interfaz de usuario gráfica

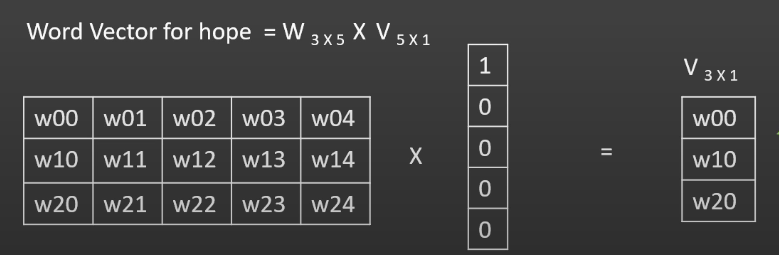
Descripción generada automáticamente

1. Se transforman las palabras por medio de un one hot vector, tomando como dimensión a los elementos únicos del corpus, siendo su entrada.Gráfico

   Descripción generada automáticamente
2. Los pesos que se generan por cada palabra en busca de acertar al target se guardan.



1. Y con los pesos podemos generar los Word vectors, utilizando los pesos de la red y sus respectivos One Hot vectors.



1. Una vez teniendo los Word vectors y la matriz de pesos, se puede obtener vectores con la probabilidad de ser vecino de la palabra en cuestión.

Diagrama

Descripción generada automáticamente

* 1. Esto por medio de softmax, aplicando a cada palabra y así generando una probabilidad por todos los vectores.

Interfaz de usuario gráfica

Descripción generada automáticamente

## ¿Para qué funciona?

Todo este proceso sirve para dada una palabra y el contexto que se recabo del texto, se puede dar una múltiple solución sobre que palabras podrían estar antecediendo o procediendo a la entrada, esto genera una contextualización de las palabras (semántica).

## Aplicación

Una sencilla sería un sistema de autocompletado mucho más eficiente y adaptativo al usuario, esto recabando la información que el mismo escribe y de los sitios que frecuenta.

En cuanto a la traducción de textos mejoraría bastante al evitar la traducción directa de la palabra y dando mayor sentido a la traducción completa, además de mayor cohesión.