

## 7.- Lenguaje Natural

viernes, 15 de marzo de 2024 06:57 p. m.

### **\*\*Técnicas de normalización de textos\*\***

- Reemplazar contracciones o abreviaciones
- Reemplazar acentos
- Reemplazo de emojis por texto
- Reemplazo repetición de caracteres
- Reemplazo Intercambio de sonidos por números
- Reemplazo Intercambio de letras por números
- Reemplazo Palabras juntas

### **Word Embedding**

- Si considera el contexto de las palabras.
- Se encuentra en medio del input y hidden de una red neuronal

\*Esquema general de una red neuronal:

| Input |---->| Embedding |---->| Hidden |---->| Output |

El Embedding es una serie de redes neuronales que están conectadas con la red neuronal principal.

Se genera un vector por palabra donde se interpreta la relación que tiene con las demás palabras.

Un vector de un solo token que busca la relación con las demás palabras del vocabulario del documento.

Tamaño: 300-400 x vocabulario

(Entre más texto se le meta al entrenamiento mejor se entiende el contexto)

Que pasa con las palabras delante, con las palabras detrás, para darle una relación del token con el texto.

Con embedign se utiliza el metodo de one hote encoding.

Se tiene 1 neurona para cada término del vocabulario y sus entradas son aplicadas con el metodo de one hot encoding.

El texto a procesar se genera una matriz de dimensión de palabras de contexto x vocabulario.

La obtención de los términos es una multiplicación matricial entre el método de Embedding por la matriz de dimensión de palabras de contexto x vocabulario.

El resultado será un solo vector del que se indica cual fue la palabra a buscar. En realidad en términos computacionales se busca el índice de la palabra para extraer el vector en esa posición.

**Función de activación:** Se coloca en todas y cada una de las neuronas de la capa intermedia de la red neuronal.

En Embedding se utiliza la función de la recta como función de activación:  $y = mx + b$

**Función de control de error:** Se coloca en la capa de salida de la red neuronal.

La m dentro de la función de activación se calcula con el gradiente descendente.

La b es una desviación que es única para cada conjunto de neuronas.

(N1)  $m1 \cdot x1 + bA$

(N2)  $m2 \cdot x2 + bA$

Como funciona el gradiente descendente, debe tener un límite para que no se vaya al infinito.