

Representación distribuida de palabras, frases y su composicionalidad.

¿Qué es?

Son métodos para optimizar al modelo Skip-gram, enfocándonos en mejorar el tiempo de calculo de softmax y además en búsqueda de permitir que el modelo reconozca frases que sus significados por separado no tienen relaciones, pero en el mundo si las tiene (Ney York Times).

¿Cómo funciona?

Hay varios métodos que en general se implementan para mejorar al modelo:

1. Hierarchical Softmax
 - a. Este método utiliza arboles binarios para cada nodo de salida de la red neuronal y por medio de él y los pesos obtenidos, genera la distribución de probabilidad, solo teniendo que evaluar $\text{Log}_2(W)$, en lugar de W volviendo más ligero el proceso
2. Noise Contrastive Estimation (NCE)
 - a. Este método busca aproximar el logaritmo de la probabilidad que arrojaría softmax
 - b. Busca que el modelo distinga los datos del ruido
3. Negative Sampling (NEG)
 - a. Es una simplificación de NCE, puesto que no indaga en todas las posibles palabras que pueden ser aledañas a la entrada.
 - b. Se le otorga k palabras negativas, esto es que esperamos que retorne 0 en su salida y así mejoramos la afinidad de la palabra de entrada sin modifican significativamente las k muestras.
4. Subsampling
 - a. Consiste en las palabras de alta frecuencia irlas quitando total o parcialmente, puesto que llegan a no denotar mucho contexto y aún después de millones de iteraciones no tienen cambios significativos.

¿Pará que sirve?

Todo esto sirve para reducir la carga computacional al Skip-gram y con ello poder utilizarlo para ponderar las palabras aledañas a la palabra que se da, se genera un espacio vectorial con relaciones geométricas que representan relaciones en la vida real y reconocimiento de expresiones más allá de su significado individual.

Aplicaciones

Generación de analogías, esto es a partir de una relación dada (puede ser Gato-Pescado, si queremos algo similar para Perro es probable que nos arroje para este sujeto otra palabra que tiene la misma relación). Otra aplicación más directas son intérpretes o traductores mejores y predictores de texto mucho más acertados y adaptativos al usuario.