

Resumen papers

Lautaro Estienne

May 26, 2020

1 2011

En este paper tratan de aprender word embeddings que sirvan para usar en tareas de sentimientos. Cuando uno aprende word embeddings en general muchas veces no se logran capturar sentidos específicos de la palabra, que son los que me interesan para la aplicación (justamente, por ejemplo, la connotación sentimental de una palabra).

El aprendizaje de estos vectores se hace a maximizando una función de costo que consiste de la suma de dos términos:

- El primero, es el "costo de las palabras del documento". Es decir, se define una función de costo que depende de las palabras en un documento y se maximiza esa función. Los parámetros aprendidos son los word embeddings. Este es un método totalmente no supervisado. Simplemente se está estimando una distribución $P(d)$ a partir de las muestras d_1, \dots, d_N .
- El segundo es el costo supervisado de la clasificación de sentimientos (igual que logistic regression).