

≡ Item Navigation

Vocabulary & Feature Extraction

Given a tweet, or some text, you can represent it as a vector of dimension V , where V corresponds to your vocabulary size. If you had the tweet "I am happy because I am learning NLP", then you would put a 1 in the corresponding index for any word in the tweet, and a 0 otherwise.

I am happy because I am learning NLP



$$\begin{matrix} [\theta_0, \theta_1, \theta_2, \dots, \theta_n] \\ n = |V| \end{matrix} \longrightarrow \begin{cases} 1. \text{ Large training time} \\ 2. \text{ Large prediction time} \end{cases}$$

As you can see, as V gets larger, the vector becomes more sparse. Furthermore, we end up having many more features and end up training θV parameters. This could result in larger training time, and large prediction time.

✓ Completed

Go to next item



Like



Dislike



Report an issue