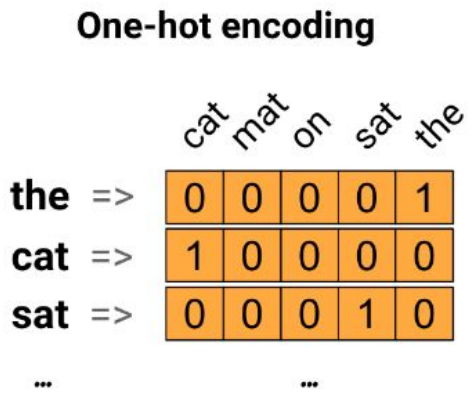
WORD EMBEDDINGS:

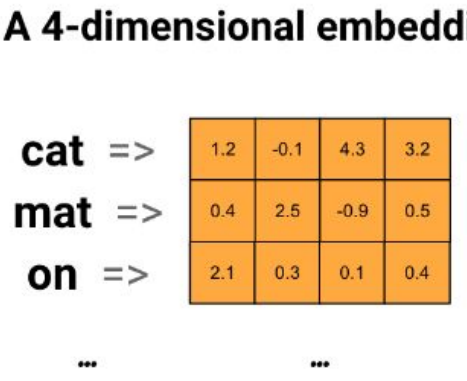
- Strategize to convert text to numbers.
- Machine learning models feed on vectors.Methods to convert text to numeric vectors.
    - **<u>One-hot-encoding:</u>**



        - Inefficient method coz sparse vectors.

    - **<u>Encode each word with a unique number:</u>**
        - Efficient because dense vector now. Like [5,3,4,1,0,2]
        - **Inefficient**
            - Integer encoding is arbitrary.
            - No meaningful relation for a model to interpret.

    - **<u>Word Embeddings:</u>**
        - Efficient dense representation, similar words have similar encoding.
        - Dense vector of floating point value, which are trainable.
        - Higher the dimension of embedding, better the embeddings, but require larger datasets.



Configure the dataset for performance