How did the result change with the various modifications? What properties of the vectors do you expect to observe when setting left (right) window size to 0? Did you manage to observe these properties?

Change the dimensionality of the vectors to 10 with 8 non-zero elements (try different dimensionalities and number of non-zero elements):

- Dimensionality refers to the number of elements in each vector. Higher dimensionality allows NN to capture more distinctions betweens words and can read to better " word embeddings" however computationally expensive. Lower simplifies it. Can read to colissions?

- Having fewer non-zero makes the vectors sparse which retains computational efficiency. Too sparse however makes the embeddings words

- Increasing window size allows more words to contribute to the context vector of each target words

- Symettrical or unsymmterical: Did not make that much of a difference. Larger "right side" makes it focus on "future" words while on the left side makes it focus on "past words". Language dependent.