**Cosine similarity Metrics:**

**Why are we using Cosine similarity Metrics:**

1: The key reason for using cosine similarity lies in its ability to evaluate how similar two vectors are, based on the angle between them in a multi-dimensional space, regardless of their magnitude.

2: In our project, we must find out the **Semantic similarity** between two texts. Therefore, this metrics will help us determine how similar two pieces of text are semantically.

**Scale used for measuring Cosine similarity Metrics:**

The orientation of two vectors matters the most in Cosine similarity metrics irrespective of the magnitudes of two vectors. Therefore, the direction of two vectors defines how similar they are. For example, two documents might have different lengths but could still contain similar content.

**Cosine similarity formula: A mathematical equation with black text

Description automatically generated**

**Interpretation of Results:**

The output score lies between -1 and 1.

If:

Score = 1, Vectors are identical (this mean they are in the same direction and angle between them is 0) as shown in fig below.

**A graph of a function

Description automatically generated with medium confidence**

Score = 0, Vectors are orthogonal (this means they are 90 degrees apart) as shown in fig below.

**A graph of angles and angles

Description automatically generated**

Score = -1, Vector are opposite to each other (this means they are completely dissimilar) as shown in fig below.

A graph of angles and angles

Description automatically generated

**Reference:** [**https://www.learndatasci.com/glossary/cosine-similarity/**](https://www.learndatasci.com/glossary/cosine-similarity/)