**How to Do Speech Recognition with a Dynamic Time Warping Algorithm**

Speech recognition is the process of analyzing audio signals to identify the words spoken by the speaker. Speech recognition has been part of our lives since 1952. We all use this technology in many areas of our daily life. Siri, a digital assistant, is an example of this technology.

The dynamic time warping (DTW) algorithm, invented by Soviet researchers in 1970, was used in speech recognition. Over time, new algorithms replaced it, but it is still a popular technique.

The dynamic time warping algorithm is a dynamic programming algorithm and a very popular technique in speech recognition. Why is the DTW algorithm a suitable method for speech recognition? Today we will see how this technique is used in speech recognition. Let’s learn this together.

Focusing on the solution without fully understanding the problem is one of the biggest mistakes we can make. Let’s first understand the problem. One of the problems in speech recognition is speed.

Diagram

Description automatically generated

DTW creates a shift in time and maps each element in the series to the closest element in the other series. In other words, DTW finds the optimum distance(the optimum distance here is the shortest distance) between the elements while doing this mapping. So a time shift will be created.

A picture containing diagram

Description automatically generated

These distances between points are stored in a table. This is called memorization. Then the shortest paths are added, and this is our similarity measure between two time series. Remember, we are trying to detect the similarity between the two signals.

Chart, scatter chart

Description automatically generated

While all this is happening, a path called the warping path is created. Time is shifted according to this path, so the two series reach the same time levels. As the warping path becomes smaller, the similarity between the two-time series increases. The warping path occurs according to some rules. The warping function represents these rules. The warping function is applied to both series. Time alignment takes place in this way. This function contains some restrictions: monotonicity, continuity, boundary conditions, and warping window. Thanks to these restrictions, the paths to be tried are limited.

Let me show an example. We have two time series.



A picture containing text

Description automatically generated