ITI41820-1 23V Advanced Topics in Machine Learning

ECG Classifier with Dynamic Time Warping

Kim kristoffer Pal
03.05.2023

Contents

1 Introduction

This work will deal with the classification of ECG data with Dynamic Time warping, and it will explain the use of DTW in general and the functionality of the DTW algorithm. It will also lock into other use cases/work where the DTW algorithm has been used. This report also locks into the implementation of the DTW algorithm in the ECG classifier that is to be implemented with phyton. At least the result of the implementation would be presented and discussed.

Dynamic Time Warping (DTW) Is an algorithm to perform time series analysing ('Dynamic Time Warping', 2023). The DTW is a dynamic programming algorithm. "Dynamic programming is both a mathematical optimization method and a computer programming method" ('Dynamic Programming', 2023). In both situations, it refers to simplifying a complicated problem by dividing it down into simpler smaller chunks of problems ('Dynamic Programming', 2023).

The Way DTW work allows it to measure the similarity between two signals that are of different length or start and end at different times ('Dynamic Time Warping', 2023).

Because of the way the DTW work it could be suited to classify ECG signals, which is what would be explored and reported on in this report.

ECG data could be used to control the functionality of a heart and it can be used to identify different heart abnormalities. A normal heart has a certain characteristic of the time-series signal measured. And the abnormal time-series signal has the same. Ewen though ECG data can be used for identifying different abnormal heart conditions, this work would only lock into the classification of normal or abnormal ECG signals.

Ter is four articles that would be presented with the objectives, working principle/methodology, results/achievements, and the advantages/disadvantages of the articles.

2 Machine Learning Algorithm

The chosen algorithm for this project work is Dynamic Time Warping (DTW) Which is an algorithm for time-series analysis. It can be used on all data that can be turned into a linear sequence. As mentioned in the introduction the DTW is a dynamic programming algorithm. And dynamic programming refers to making some complicated problems easier to solve by breaking them down into smaller simple problems.

The DTW algorithm is special or noteworthy if you will because it could measure the similarity between two signals that are not of equal length or do not have the same stat time. Therefore, the DTW is capable of comparing two sequences that have a similar shape but with local variation (Guo et al., 2022).

Compared to other similar algorithms such as Euclidean Distance (ED) or Pearson correlation the main difference from DTW is the way they measure the similarity between sequences. The ED measure the straight-line distance between two points in a multidimensional space. In other words, it measures the distance between two points in two sequences based on the difference between their corresponding values in each dimension. DTW, on the other hand, calculates a distance measure that considers the temporal distance between corresponding points in the two sequences. In Figure 1 the difference between ED and DTW is illustrated.

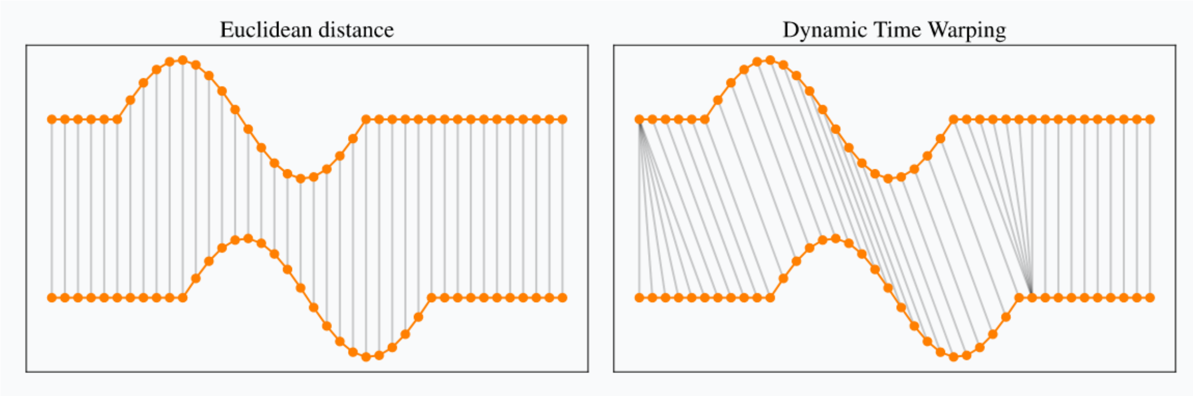

Figure 1 illustration of the difference between ED compared to DTW. By (Tavenard, 2021)

The DTW distance is the minimum cumulative distance needed to align the two sequences. The ED is commonly used as a distance metric in many machine learning algorithms. Where DTW is used to measure the similarity between two sequences that may vary in their length and speed for instance if we want to compare two people walking at different speeds and not starting at the same time, but still having the same walking pattern. In summary, ED measures the straight-line distance between two points in two sequences, and DTW measures the temporal distance between corresponding points in two sequences while allowing for time warping to align them. Then what about the Pearson correlation?

Pearson correlation measures the linear relationship between two variables by calculating the covariance of the variables divided by the product of their standard deviations. It is a measure of how much two variables are related to each other, and it ranges from -1 (perfect negative correlation) to 1 (perfect positive correlation). In the context of sequence analysis, Pearson correlation can be used to measure the similarity between two sequences by calculating the correlation between the corresponding values in each sequence. As mentioned in the comparison with the ED. The DTW on the other hand considers the temporal structure of the sequences by allowing for warping or stretching and compressing, the time axis to align the sequences. To summarise Pearson correlation measures the linear relationship between two variables, and DTW measures the temporal distance between corresponding points in two time-series data sequences while allowing for time warping to align them. In the next section w will lock at the working principle of the DTW.

2.1 Working principle of the DTW

Shortly the Idea of the DTW is to find an optimum match between two signals, this is done by calculating the cost between two sets of data(sources). The cost in the case with DTW is the sum of the absolute difference for each matched pair of indexes of two data sets. The less difference means a smaller error which again means a better match. To understand more deeply how it is working, it is necessary to understand a set of rules for the DTW:

- Each index from the first series needs to be matched with one or more indexes from the other series and vice versa.
- First indexes from the first series need to be matched with the first index in the other series, however, it does not need to be its only match.
- Last indexes from the first series need to be matched with the last index in the other series, however, it does not need to be its only match.
- Mapping of indexes from the first series to the indexes from the other series need to only increase, and vice versa, for instance, if $j > n$ is indexes from the first series, then it must not be

two indexes $n > m$ in the other series, so the index i is matched with index n and j is matched to index m, and vice versa.

The illustration in Figure 2 shows how two signals are compared and illustrate the calculated distance and matching between the two signals.

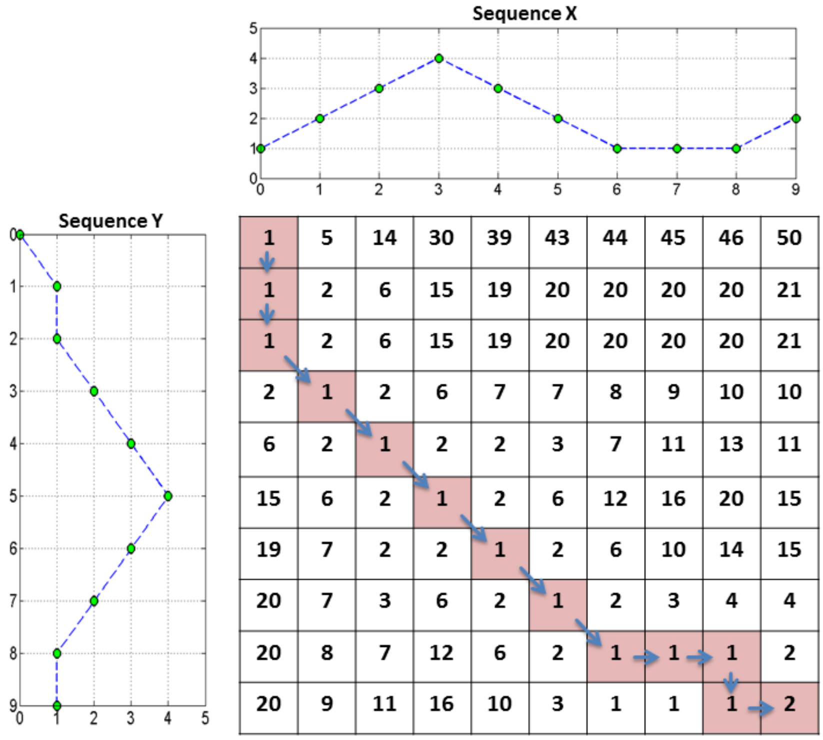

Figure 2 illustration of DTW similarity between the x series and y, of length 10. And the lowest warp path is indicated with the arrow pointing along the coloured boxes. By (Al-Jawad et al., 2012)

In the figure x above, we can see that index 0 to 2 in the y sequence is pared agents the value of index 0 in the x sequence which also illustrates the two first rules. In the end, we can also see that index 9 in each of the sequences is matched with each other. Which illustrates the third rule.

We can also explain the DTW more mathematically with formulas or equations.:

Let us say we have two sequences of data x and y where x = {x1, x2, …, xn} y = {y1, y2, …, ym} then the DTW aims to find the optimal alignment between the two sequences by minimizing the cumulative distance between corresponding elements. The cumulative distance between elements xi and yj is defined as: D(i,j) = (xi -yj)^2 The DTW algorithm constructs a two-dimensional matrix D with dimensions (n+1) x (m+1), where D(i,j) represents the cumulative distance between elements I and yj. To compute the DTW matrix D, we start by initializing D(0,0) = 0, D(i0) = infinity for 1 <= i <= n, and D(0,j) = infinity for 1 <=j <= m. This sets the first row and column of the matrix to infinity, ensuring that the algorithm starts at the first element of both sequences.

The Next step is to iterate over all the remaining elements of the matrix, computing the minimum cumulative distance between each pair of elements using the following recurrence relation:

D(i,j) = d(i,j) + min(D(i-1,j), D(I,j-1, D(i-1,j-1))

The relation above computes the cumulative distance between elements xi and yj by adding the distance between the elements to the minimum of the possible paths:

- The path from the previous element of sequence x to the current element of sequence y (i.e., D(i-1,j))

- The path form the current element of sequence x to the previous element of sequence y (i.e., D(i,j-1)).
- The pat form the previous element of sequence x to the previous element of sequence y (i.e., D(i-1,j-1)).

By recursively computing the minimum cumulative distance between all possible pairs of elements, the algorithm finds the optimal alignment between the two sequences x and y. To extract the optimal alignment path from the DTW matrix D we can start at the left upper corner (i.e., D(n,m)) and trace the lowest cumulative distance, choosing the path with the minimum cost at each step as illustrated in Figure 2. In the next section, we will investigate the advantages and disadvantages with the DTW algorithm.

2.2 Advantages and Disadvantages with DTW

This section would look at the advantages and disadvantages with the DTW algorithm.

The first advantages with the DTW are that it can align sequences with variations in time and speed. This makes it particularly useful in applications such as speech recognition, music analysis and gesture recognition where the length and speed of the input sequences may vary. Secondly, DTW is a flexible algorithm that can be customised to specific needs. As to changing the window size / Wrapping window, which constrains the available subset of the matrix. Setting a window size could slightly speed up the calculation time on the distance matrix, and in some cases, the right window size could improve classification accuracy.

Some disadvantages with the algorithm. First, if a noticeably short sequence is compared to a long signal the distance score would be low even though the signals are not similar. In such a situation the score would need to be normalised to avoid that problem. A second disadvantage is that on large datasets with long sequences, it may be impossible to calculate the distance score in a realistic amount of time due to the dime complexity. Lastly, it may not be suitable to use the DTW algorithm for data that does not have a clear pattern or structure. In the next section, we would look at some examples on practical applications where DTW is used.

2.3 Examples of practical applications where DTW is more suitable than other technique(s)/algorithm(s)

Historically, DTW is used on sound or audio clips to classify similarity in audio clips (*Understanding Dynamic Time Warping*, 2019). And that works well because of the ability of DTW to compare signal with different starting time. For instance, the two audio clips contain the word hello but in one of the recordings the word is spoken slowly almost ass “helllllloooooo” compared with “hello!” it is still the same word but the alignment of the signals might not be the same since the first one would haw more signals over a longer time period than the fast speaking, here DTW would be able to classify the tow word as the same. Where the ED algorithm would haw classified it as two different words. One last example could be wearable fitness trackers as calculating a walker’s speed and the number of steps, even if their speed varied over time. In this example let us also compare with the ED algorithm which would not haw be able to manage it because of the different starting times and walking speeds, However, the DTS algorithm is suitable to it because of its ability to detect similarities even if the start time and walking speed is different.

3 Chosen Application

The application in this study is an Electrocardiogram (ECG) data classifier to classify if ECG data is normal or abnormal. The idea is that the classifier could haw a set or collection of reference ECG data that is pre classified as normal or abnormal. Then it can take unclassified ECG data as input to be classified as normal or abnormal. The idea behind the application is that it could be used to flag the abnormal ECG data to be closer investigated by a doctor or specialist.

3.1 Challenges with the Application

Challenges associated with such application could be.

- To ensure the correctness of the classification
- Adapt it to the variability of ECG signals. Due to factors such as patient age, gender, health conditions and medication.
- The data quality can be affected by a range of external factors, such as electrode placement and motion artefacts.

3.2 Why The algorithm is suited to the selected Application

The choice of DTW to manage the application is because it is a technique that can compare two time-series data regardless of the difference in their length and shape. And since that is the case with ECG signals that they can vary significantly in shape and length due to differences in heart conditions. The DTW algorithm can capture the variation in the time series data and still providing accurate classification results it is deemed to be suitable, to the ECG classifier application.

3.3 Other alternative algorithms

Other alternative algorithm could be machine learning-based techniques such as support vector machines , artificial neural network, and decision Tree Induction (Turnip et al., 2018), (Sánchez & Cervera, 2019), (Kaur & Singla, 2016). The author of the paper on support vector machine came to haw reach an classification accuracy of 88.49% (Turnip et al., 2018). And the authors of the paper on using Artificial Neural Networks claim to reach an average accuracy at 97.6% (Sánchez & Cervera, 2019). The authors of the article using Decision Tre incubation claims to haw reached an average accuracy of 100% on a standard dataset (Kaur & Singla, 2016).

I believe that my selected DTW algorithm is better than the alternatives because the machine leering-based techniques may require a large amount of labelled data to train a model, which may be difficult to obtain in the case of medical applications such as ECG classification.

4 Related work / Literature

In this section we would lock into the for selected relevant articles that uses the DTW algorithm in an application. And we would lock into the main goal, methodology and advantages/disadvantages with the selected articles. All the four selected articles are selected based on that they all are using DTW algorithm in one way or another.

- Dynamic time warping using graph similarity guided symplectic geometry mode decomposition to detect bearing faults.
- Dynamic Time Wrapper Based Local Predictor for Wind Speed Prediction
- Assessment of Fire Regimes and Post-Fire Evolution of Burned Areas with the Dynamic Time Warping Method on Time Series of Satellite Images—Setting the Methodological Framework in the Peloponnese, Greece
- Intelligent Diagnosis of Abnormal Charging for Electric Bicycles Based on Improved Dynamic Time Warping

This chapter would be organised in the same order as the order of the articles listed above. With subsections with the same title as the papers.

4.1 Dynamic time warping using graph similarity guided symplectic geometry mode decomposition to detect bearing faults

4.1.1 Objectives

The objective of this article is to propose a new method to detecting bearing faults in machinery using DTW and symplectic geometry mode decomposition guided by graph similarity(source). And bay that overcomes limitations of traditional vibration analysis methods, that are not able to accurately detect bearing faults under varying operating conditions (Guo et al., 2022).

4.1.2 Methodology

The methodology in the article on bearing faults is divided into four main steps, first collecting vibration signals of the different health conditions on bearings (Guo et al., 2022). Second, they decompose the raw signal into numerous components which are defined as "symplectic geometric components (SGCs)" (Guo et al., 2022). Thirdly selecting of an SGC that contains sufficient fault information is selected by calculating the Graph similarity values between all SGCs and raw vibration signal (Guo et al., 2022). Then the SGCS with the lowest GS score to be the optimal SGC(Guo et al., 2022). The last step is to use the DTW algorithm to classify the test sample based on the fault type belonging to a template signal with the fault types(Guo et al., 2022). Then based on the lowest DTW distance they determine the fault it is belongs to (Guo et al., 2022).

4.1.3 Results

To evaluate the results, they used experimental data from a veering test rig and compared the results with traditional vibration analysis methods. And the final results show that the method used demonstrated better results and more correct fault classification compared to traditional methods in varying operating conditions (Guo et al., 2022). Therefore, they conclude that the presented method shows potential use in a real word bearing fault diagnosis.

4.1.4 Advantages and Disadvantages

The advantages and disadvantages of this article are presented in a list.

- Advantages
    - The proposed method could accurately detect bearing faults in varying operating conditions.
    - The DTW and symplectic geometry mode decomposition allows for the effective detect bearing faults.
- Disadvantages
    - The proposed method may require complex computation which could make it less suitable for real time monitoring.

In the next section, we will lock into the second selected article.

4.2 Dynamic Time Wrapper Based Local Predictor for Wind Speed Prediction

4.2.1 Objectives

The objective of this article is to suggest a novel wind speed prediction model that takes advantage of the DTW algorithm and local prediction approach to improve the accuracy and trustworthiness of wind sped predictions (Jiang et al., 2022). In order to verify their results it is to be evaluated with other commonly used prediction models for wind speed (Jiang et al., 2022).

4.2.2 Methodology

The methodology in the wind sped prediction article is divided into three steps. First the similarity between historical wind speed and the current data using DTW algorithm (Jiang et al., 2022). Secondly, a local prediction approach is used to estimate the wind speed in the next time step. Lastly, the model is evaluated with collected real world wind speed data (Jiang et al., 2022).

The data set they used in the project was the entire data of DBWS which has 309986 samples, and they used a train test data ratio of 7:3 (Jiang et al., 2022). They also used DTW to increase the quality of the training set, bay calculated the DTW distance between all the train and test samples and then sorted them (Jiang et al., 2022). Then they used 30% of the training samples with the lowest DTW distances as improved training data (Jiang et al., 2022).

4.2.3 Results

The results of the study showed that the proposed wind sped prediction model RF model with DTW based local predictor (RF-DTWLP) showed a reduced mean square error of 75% compared to the RF method and the mean square error is reduced by 50% compared to the RF method (Jiang et al., 2022).

4.2.4 Advantages and Disadvantages

The advantages and disadvantages in the article about the wind sped prediction is to be presented in a list underneath.

- Advantages
    - The proposed wind sped prediction model gives improved accuracy and trustworthiness in comparison to other commonly used models.
    - The adaption of DTW algorithm and local prediction gives the model the ability to better capture patterns and variation in wind speed data.
    - The results are displayed in a table.
- Disadvantages
    - Lack of detailed methodology description
    - The presented solution might require more computational resources and time compared to simpler models.
    - The model's performance may be affected by the quality and availability of the historical wind speed data used in the training.

In the next section, we would lock into the third selected article.

4.3 Assessment of Fire Regimes and Post-Fire Evolution of Burned Areas with the Dynamic Time Warping Method on Time Series of Satellite Images—Setting the Methodological Framework in the Peloponnese, Greece

4.3.1 Objectives

The objectives of the article on fire area evaluation are to demonstrate the use of DTW algorithm to asses fire regimes and post-fire evaluation of burned areas with time series satellite images in Peloponnese Greece (Koutsias et al., 2022). In the article, the aim is to provide a methodological framework to compare different images and analyse the temporal changes in vegetation cover after a fire event using DTW (Koutsias et al., 2022).

4.3.2 Methodology

The methodology in the presented article about evaluation firer affected areas is divided into three steps. The first step is to use Landsat timeseries images covering a 32-year period from 1984 to 2016 of burned and non-burned areas (Koutsias et al., 2022). The time series satellite images were then pared up with pre fire images and post-fire images. Secondly, a vegetation index is calculated based on the satellite images (Koutsias et al., 2022). Lastly, the DTW method is applied to compare different images and analyse the temporal changes in vegetation cover (Koutsias et al., 2022). The programming language used in this study was R and they used the dtw package available there (Koutsias et al., 2022).

4.3.3 Results

The results in this article show that the DTW method was effective in detecting and analysing the temporal changes in vegetation coverage after a fire event (Koutsias et al., 2022). And they were able to detect and assess changes in the landscapes as changes in vegetation after a fire incident (Koutsias et al., 2022).

4.3.4 Advantages and Disadvantages

The advantages and disadvantages of the article will be presented in a list underneath.

- Advantages
    - The result in the article is illustrated with figures with plots of the results and bar charts.
    - The satellite data used in the study is Freely available.
    - The results in the article were promising.
- Disadvantages
    - The disadvantage of this approach is that the method can be computationally intensive.

In next section, we would lock into the fourth and last selected article.

4.4 Intelligent Diagnosis of Abnormal Charging for Electric Bicycles Based on Improved Dynamic Time Warping

4.4.1 Objectives

The objective in the article about abnormal charging diagnosis is to propose a method for intelligent diagnosis of abnormal charging in electric bicycles (Shuai et al., 2023). And this is to be obtained by using an improved version of the DTW algorithm (Shuai et al., 2023). Bay the work the goal is to improve the safety and reliability of the electric bicycles and decrease the maintenance cost (Shuai et al., 2023).

4.4.2 Methodology

The mythology in the article on abnormal charging diagnosis is divided into two steps. The first step is to take a sequence with current and remove noises from it by what is described as a median filter (Shuai et al., 2023). Second and lastly they combine the longest Similar subsequence (LSS) algorithm and DTW algorithm to what they defined as the improved version of the DTW algorithm or LSS-DTW (Shuai et al., 2023). The improved algorithm is dan given collected data from charged bicycles. When this data is faded to the algorithm they get a number value out that they can categorize if the condition and state are normal or not (Shuai et al., 2023).

4.4.3 Results

The experimental results in the study on abnormal bicycle charging show that the LSS-DTW effectively can detect abnormal charging patterns and diagnose the cause of the problem (Shuai et al., 2023). Therefore it is concluded that the proposed method has the potential to improve the safety and reliability of electric bicycles and reduce maintenance costs (Shuai et al., 2023).

4.4.4 Advantages and Disadvantages

The advantages and disadvantages are presented in a list.

- Advantages
    - The method’s ability to accurately diagnose abnormal charging patterns.
    - The method can reduce maintenance costs.
    - The articles have bar charts where the result of the method is compared to multiple different methods.

- Disadvantages
    - The proposed method may require specialized equipment and expertise to implement, this could increase the cost and complexity of maintenance.

In the next section, I will present the method in the ECG data classifier.

5 Implementation and Results Analysis

In this section the implementation of the DTW algorithm is explained and the result is compared whit other similar work. We will first lock at the data set than on the DTW algorithm development and dive into the results and comparison whit a similar implementation. The entire code and data set used is available in git hub: https://github.com/kimpal/DTW_ECG_Classifier

5.1 Data Set

The data set used is found on Kaggle.com (*ECG Classification Using FastDTW*, n.d.) and it is a simplified version of the ECG5000 Data set form (*Time Series Classification Website*, n.d.) it contains labelled data where abnormal data is labelled with 0 and normal data is labelled with 1. The columns contain the time series data on ECG, and it is containing 140 data points. The rows correspond to the number of individual measurements which is equal to 4997 Rows in total. The first step that was performed on the data set was to split it into train and test data. In this case, it is some of the train data that is supposed to be the reference data in the DTW algorithm. And the labels are also split out from the data since we would not feed that into the DTW algorithm. The test size is set to 20% of the total data. The split is with a random_state = 0 to ensure the same data for every run. Just to make the reproduction of the results easier. Then the normal and abnormal data are separated to be able to know which reference data that is normal or abnormal. And I also make it possible to evaluate the number of correct versus wrong predictions easier. After all the splitting we had in total four data sets normal test normal train and abnormal train and abnormal test.

5.2 The DTW algorithm development

The first step was to randomly pic reference data to use in the DTW algorithm from the abnormal and normal train data this was done with the random module in Python where the start and end range could be chosen the start and end range were set to the total size of the abnormal test data and the normal test data. This was run three times to select in total six reference values 3 normal and 3 abnormal. It was also tested whit only two normal and two abnormal reference values as we will show in the results. When the index for the reference data was chosen the reference data was, please in a python dictionary named ECGs.

To make the dtw algorithm work first a function named dtw was defined which takes three variables x, y, and window. where x is the test sequence, y is the ECGs reference sequences and window is the variable to the window size. The following steps in the function will be listed under neat in a numbered list:

1. The length of the input signal x and y is set to the variable n and m, respectively.
2. Then the window variable is set to the maximum value between the input window argument and the absolute difference between the lengths of the two time series.
3. Initializing of a matrix of shape (n+1, m+1) is filled wit infinity values. This matrix is to be used to store the DTW distances between the two time series.
4. Then the first value at the top left corner of the matrix is set to 0 since the distance between two empty time series is 0.
5. Then a loop iterates over the values in the first time series, starting at index 1 and ending at index n.
6. Then a nested loop iterates over the values in the second time series, starting at index max(1, i-window) and ending at index min(m+1, i+window). The max and min functions ensure that the

loop only considers values that are within the window distance of the current values in the first time series.
7. Then the absolute difference between the current values in the two-time series is computed.
8. Then the last_min variable in the code is set to the minimum value in a 2x2 submatrix of dtw_matrix containing the previous minimum distance values. The i-1:i and j-1:j+1 indexing ensures that only the values in the previous row and the current column are considered.
9. Then the value in the current cell of dtw_matrix is set to the sum of the current cost and the previous minimum distance.
10. Then the distance value stored in the bottom-right cell of dtw_matrix, which represents the DTW distance between the two time series is returned.

The above-described function is then called later when the actual classification with data is to be run. In my case, I also call this function in the code that is checking different window sizes. Which is a process I now will describe:

To identify the suitable window size on the DTW Algorithm the algorithm was implemented with an array of window sizes 1, 2, 3, 4, 5 and 10 to systematically evaluate different window sizes that looped through the data and classified normal test data with all the different window sizes. In my test with the window size, I found no difference in accuracy between the tested window size but to gain efficiency the window size was set to 1 when I ran the actual prediction. In DTW algorithms it is known that a smaller window size will increase the run time. Then let us dive into how the actual classification was run and what output values we set in the next paragraph.

The code implementation to evaluate the classification method was duplicated so it is splatted up so that the classification on abnormal and normal ECG gives a separate accuracy score and total execution time. This allowed for an easier calculation on the accuracy.

The validation was a simple implementation since the test data was separated into normal and abnormal so than the number of correct predictions was only divided by the total number of rows in the test data and the product was the accuracy, and total execution time in seconds.

When it comes to the prediction part is essential to calculate the distance score on each of the test data against the reference data than the reference data that has the lowest distance score against the test data. This calculation is done for each of the test data. In other words, the lowest distance score is the measurement that decides what each test data sample is classified as.

5.3 Results and comparisons

The window size that was done with a prediction on the normal test data and the results as briefly mentioned showed no difference in accuracy. but the differences in time it takes on different window size is different the results on different window sizes are displayed under neat:

```
Testing window size: 1
Accuracy: 0.9709897610921502
Run time for window size 1: 4.975134372711182 seconds
Testing window size: 2
Accuracy: 0.9709897610921502
Run time for window size 2: 7.14889121055603 seconds
Testing window size: 3
Accuracy: 0.9709897610921502
Run time for window size 3: 11.534061670303345 seconds
Testing window size: 4
Accuracy: 0.9709897610921502
Run time for window size 4: 14.63205623626709 seconds
Testing window size: 5
Accuracy: 0.9709897610921502
```

```
Run time for window size 5: 19.209710836410522 seconds
Testing window size: 10
Accuracy: 0.9709897610921502
Run time for window size 10: 36.18400812149048 seconds
```

It is clear in the results of the test on window sizes that the window size of 1 has the lowest run time as expected. When it comes to the result on the accuracy of window sizes there is no difference. The next test result to be presented Is the accuracy result on the classification of abnormal and normal ECG data when run on a window size of 1. Accuracy on the normal ECG classification with 4 reference ECG is:

```
Accuracy: 0.9709897610921502
Total execution time: 4.06786584854126 seconds
total test data: 586

0,9709897610921502
```

From the result, we can see that the accuracy is 97.09% on a total test set of 586 normal test data sequences.

The results on the abnormal ECG classification with 4 reference ECG is:

```
Accuracy: 0.9130434782608695
Total execution time: 2.5279593467712402 seconds
total test data: 414
```

here we can observe that we got an accuracy of 91.30% on a total test set of 414 abnormal test data sequences.

To try to increase the accuracy it was also done a run whit in total 6 reference ECGs that showed some improvement in the accuracy. Accuracy on the normal ECG classification with 6 reference ECG is:

```
Accuracy: 0.9641638225255973
Total execution time: 4.690285921096802 seconds
total test data: 586
```

Her we can observe that we got an accuracy of 96.41% on a total test set of 586 normal test data.

The results on the abnormal ECG classification with 6 reference ECG is:

```
Accuracy: 0.9758454106280193
Total execution time: 2.752830982208252 seconds
total test data: 414
```

Her we con observe that we got an accuracy of 97.58% on a total test set of 414 abnormal test data.

On the test run whit 6 reference data we can observe that the accuracy on the normal test data is a bit lower but the accuracy on the abnormal data has increased so the score lock to be more evenly between the normal and the abnormal classification. And that is something we want to se. Because than the total accuracy is higher. How ewer whit no comparison to other results it is not easy to say how good or bad the result is.

To comparison the results I found an article where they haw classified ECG data whit an combination of an Deep Neural network and DTW (Ahmadi-Mobarakeh & Mohammadzade, 2021). Their results showed an accuracy of 92.3% on a similar data set. The only difference on the data sets is that as mentioned on the section on Data Set is that I used a simplified version of it where each abnormal data has been labelled 0 and normal data has been labelled 1. The reasons why I got a higher accuracy than to the compared result is likely to be because they are managing 5 classes of the data but in my application the data is only labelled in two different classes normal and abnormal. In addition, the result

I get depends on the reference data selected randomly from al the Train data set. However, it is not easy to exactly pinpoint the reasons since they are using a Neural network where they firs apply the DTW and then a kernel layer and finish whit a fully connected network (FCN) so in their model there is a higher number of variables that can affect their results. They also compare their results whit other techniques and al the compared techniques. On the eECG500 Data set is giving better result than their results. How ewer They haw also tested their method whit a smaller data set ECG200 that only haw tow classes normal and a specific type of abnormal reading, when they get better results than the methods, they compared agents. The results they got is displayed in Figure 3.

Method		**Datasets**	
		ECG200	*ECG5000*
Our method		**91.0 %**	92.3 %
Deep methods	Inception Time [4]	89.6 %	94.2 %
	ResNet [5]	83.9 %	94.1 %
Non deep methods	HIVE-COTE [1]	85.8 %	94.5 %
	BOSS [14]	87.8 %	94.0 %
	ROCKET [15]	89.9 %	**94.7 %**
	WEASEL [16]	85.9 %	94.5 %

Figure 3 the results the compared paper achieved and their compared methods.

The fact that they got better results than their compared method on the smaller data set is something I think indicates that the different number of classes int the data set influences the results.

Overall, the obtained Result on the DTW classifier was positively However it might still need some improvement in the results to be applied as an application to be used in a hospital as we will lock into in discussion section.

6 Discussion

Aw we already have discussed the comparison of the results in the end of the Results and comparisons section, the discussion part in this chapter will be quite short and focuses more on the practical aspect of using the method in the practical application.

The reference data used in the DTW classifier can vary a lot so in future the reference values should be selected by an expert in ECG signals so it could be a better average representation of normal and abnormal ECG signals. As this can affect the accuracy of the prediction. Another point is that it might not be suitable to implement the solution even though it archives an accuracy score of 97.58% abnormal data and 96.41% on the normal data. Because it is critical to classify an abnormal ECG as normal science the consequences could be fatal and in the worst case could lead to death. However, the results from a scientific perspective in the DTW algorithm are promising.

7 Conclusion

Shortly the results on the ECG classifier showed an accuracy of 97.58% on classification on abnormal ECG data and an accuracy of 96.41% on et classification on the normal data. Suggestions to future work on the DTW classifier: Since the reference data can vary a lot as mentioned in the discussion section it might be wise to haw an expert select the reference data. In future work it could be interesting to expand it to classify specific types of abnormal ECG data.

8 References

Ahmadi-Mobarakeh, M., & Mohammadzade, H. (2021). ECG Classification Using DTW - Based Learnable Kernels in Deep Neural Networks. *2021 28th National and 6th International Iranian Conference on Biomedical Engineering (ICBME)*, 90–94. https://doi.org/10.1109/ICBME54433.2021.9750288

Dynamic programming. (2023). In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=Dynamic_programming&oldid=1148825637

Dynamic time warping. (2023). In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=Dynamic_time_warping&oldid=1150247571

ECG Classification using FastDTW. (n.d.). Retrieved 3 May 2023, from https://kaggle.com/code/aavigan/ecg-classification-using-fastdtw

Guo, J., Si, Z., Liu, Y., Li, J., Li, Y., & Xiang, J. (2022). Dynamic time warping using graph similarity guided symplectic geometry mode decomposition to detect bearing faults. *Reliability Engineering & System Safety*, *224*, 108533. https://doi.org/10.1016/j.ress.2022.108533

Jiang, J., He, C., & Yang, Q. (2022). Dynamic Time Wrapper Based Local Predictor for Wind Speed Prediction. *IEEJ Transactions on Electrical and Electronic Engineering*, *17*(1), 148–150. https://doi.org/10.1002/tee.23497

Kaur, B., & Singla, S. (2016). ECG analysis with signal classification using Decision Tree Induction (DTI). *Proceedings of the International Conference on Advances in Information Communication Technology & Computing*, 1–6. https://doi.org/10.1145/2979779.2979874

Koutsias, N., Karamitsou, A., Nioti, F., & Coutelieris, F. (2022). Assessment of Fire Regimes and Post-Fire Evolution of Burned Areas with the Dynamic Time Warping Method on Time Series of Satellite Images—Setting the Methodological Framework in the Peloponnese, Greece. *Remote Sensing*, *14*(20), Article 20. https://doi.org/10.3390/rs14205237

Sánchez, F. A. R., & Cervera, J. A. G. (2019). ECG Classification Using Artificial Neural Networks. *Journal of Physics: Conference Series*, *1221*(1), 012062. https://doi.org/10.1088/1742-6596/1221/1/012062

Shuai, C., Sun, Y., Zhang, X., Yang, F., Ouyang, X., & Chen, Z. (2023). Intelligent Diagnosis of Abnormal Charging for Electric Bicycles Based on Improved Dynamic Time Warping. *IEEE Transactions on Industrial Electronics*, *70*(7), 7280–7289. https://doi.org/10.1109/TIE.2022.3206702

Time Series Classification Website. (n.d.). Retrieved 3 May 2023, from http://www.timeseriesclassification.com/description.php?Dataset=ECG5000

Turnip, A., Rizqywan, M. I., Kusumandari, D. E., Turnip, M., & Sihombing, P. (2018). Classification of ECG signal with Support Vector Machine Method for Arrhythmia Detection. *Journal of Physics: Conference Series*, *970*(1), 012012. https://doi.org/10.1088/1742-6596/970/1/012012

Understanding Dynamic Time Warping. (2019, April 30). Databricks. https://www.databricks.com/blog/2019/04/30/understanding-dynamic-time-warping.html

9 List of figures