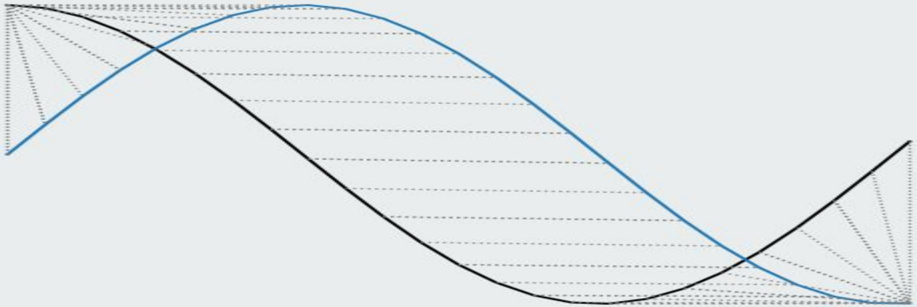
Warping Time: Using Dynamic Time Warping to Improve Machine Learning Outcomes



Kim Kraunz
ODSC West - November 2, 2023

MIMOTO

About Me

- Background in molecular biology and biostatistics
- Past 7+ years have worked as an internal data scientist and consultant

Currently

- Lead Data Scientist at Mimoto
- At Mimoto we using signal processing to identify whether someone using an account is unknown or known, and if known, who they are.

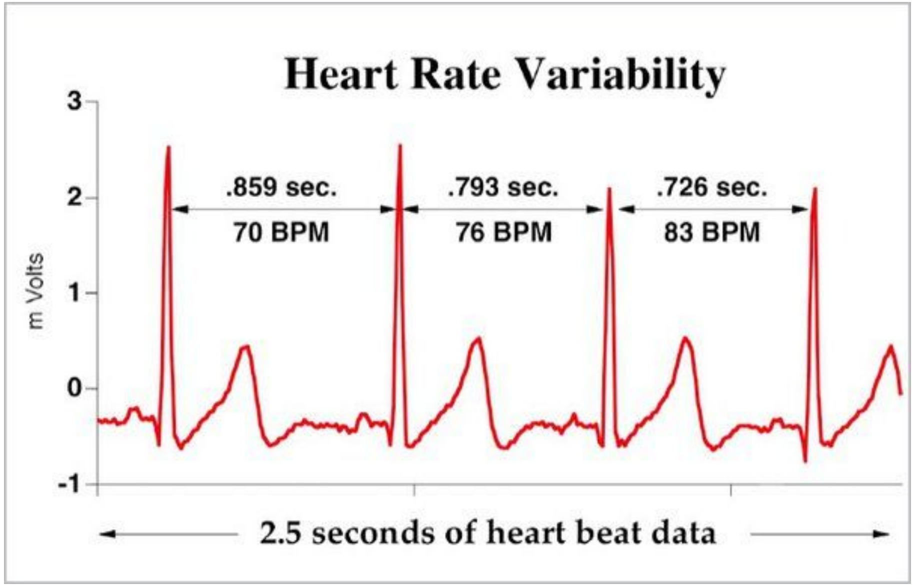
Take-Aways

1. Understand what dynamic time warping is
2. General understanding of the available packages (python and R)
3. Examples of how it boosts outcomes when combined with machine learning techniques

REPO: https://github.com/kimkraunz/odsc_west_2023

MIMOTO

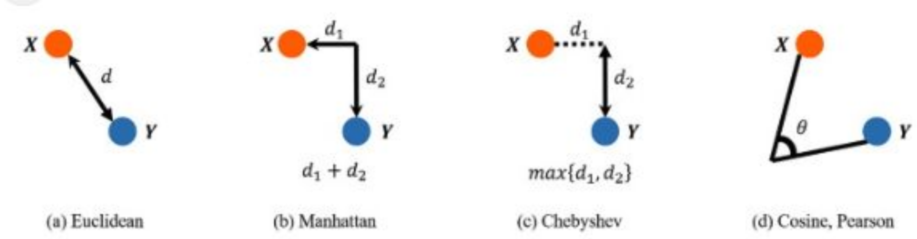
What if the time series aren't in sync?



Heart beats look very similar but have different lengths.

Goal: Measure the Similarities between Time Series

In a simple world we would calculate the distance using one of many distance measures



Ao Li, Cheng Fan, Fu Xiao, Zhijie Chen, Distance measures in building informatics: An in-depth assessment through typical tasks in building energy management, Energy and Buildings, Volume 258, 2022, 111817

Traditional methods wouldn't work to align them

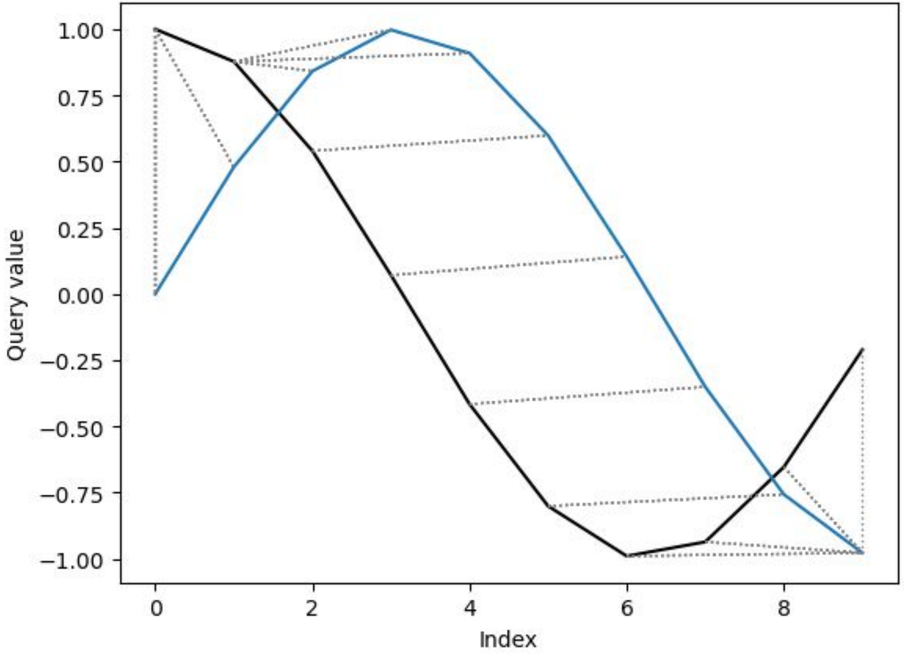
Dynamic Time Warping!

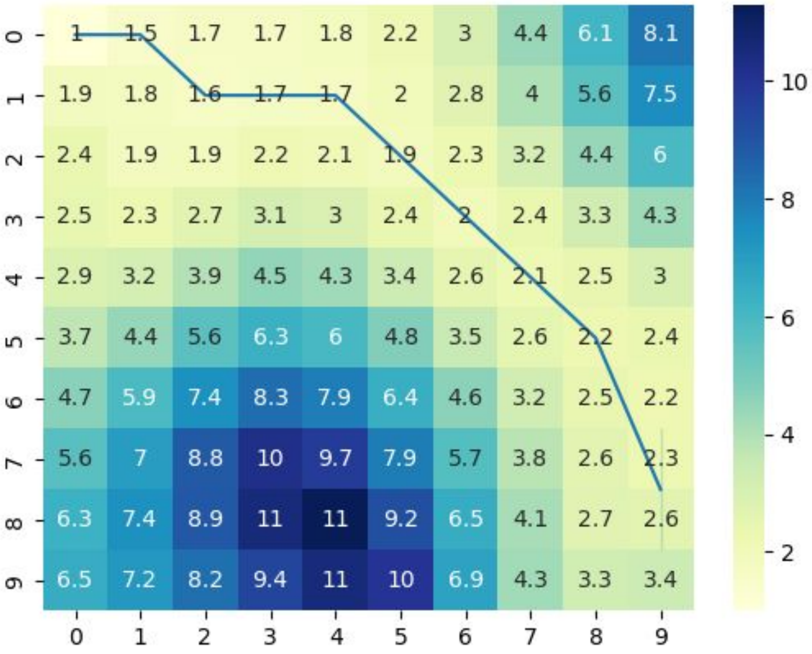
MIMOTO

Time Warping Calculation

1. Divide the two series into equal points.
2. Calculate the distance (usually euclidean) between the first point in the first series and every point in the second series. Store the minimum distance calculated. (this is the 'time warp' stage)
3. Move to the second point and repeat 2. Move step by step along points and repeat 2 till all points are exhausted.
4. Repeat 2 and 3 but with the second series as a reference point.
5. Add up all the minimum distances that were stored and this is a true measure of similarity between the two series.

Example


Query value
Index
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
0
2
4
6
8


0
1
2
3
4
5
6
7
8
9
10
8
6
4
2

Dynamic Time Warping Packages

Python

1. dtw-python
2. dtaidistance.dtw
3. TSLearn.metrics.dtw
4. pyts.metrics.dtw
5. FastDTW

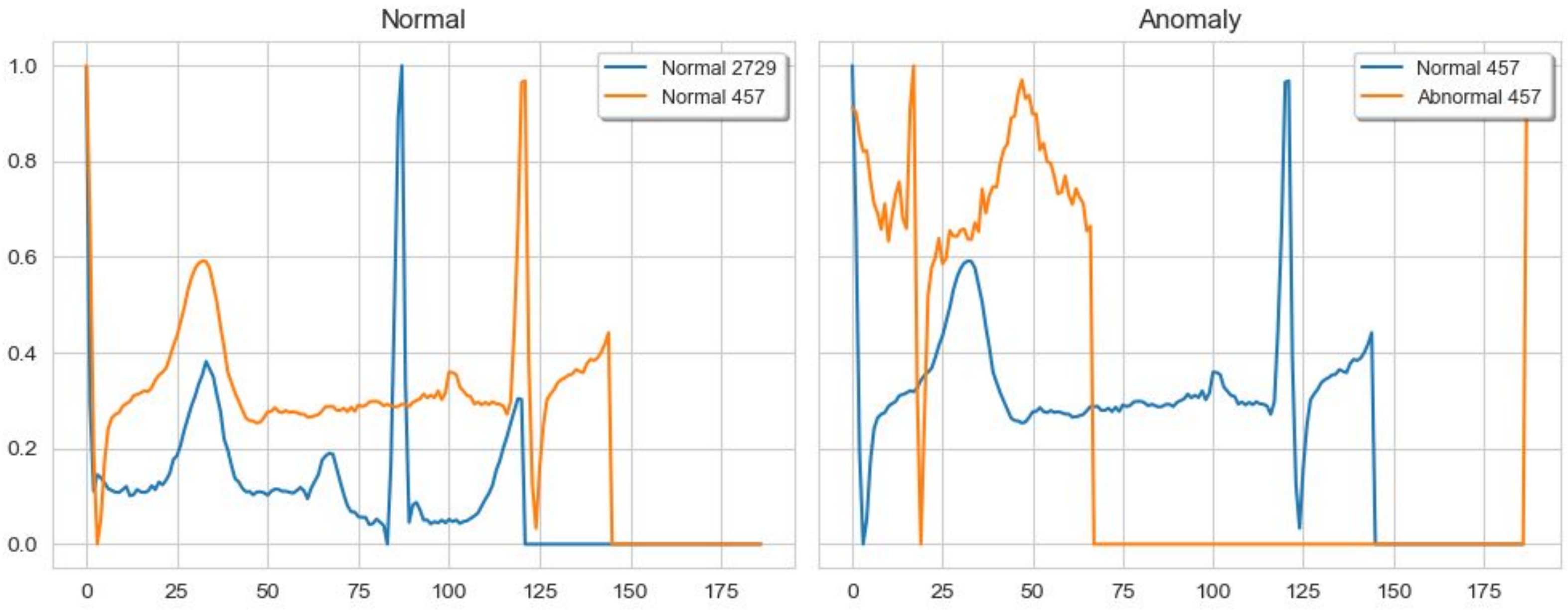
R

1. dtw
2. IncDTW

https://forecastegy.com/posts/dynamic-time-warping-dtw-libraries-python-examples/

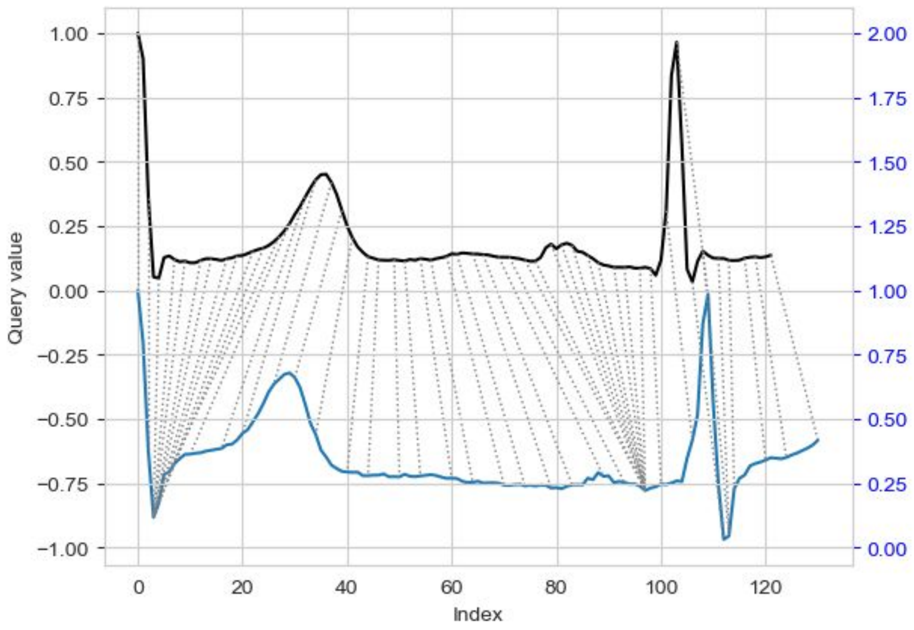
Use Cases for Dynamic Time Warping

Example - [ECG]



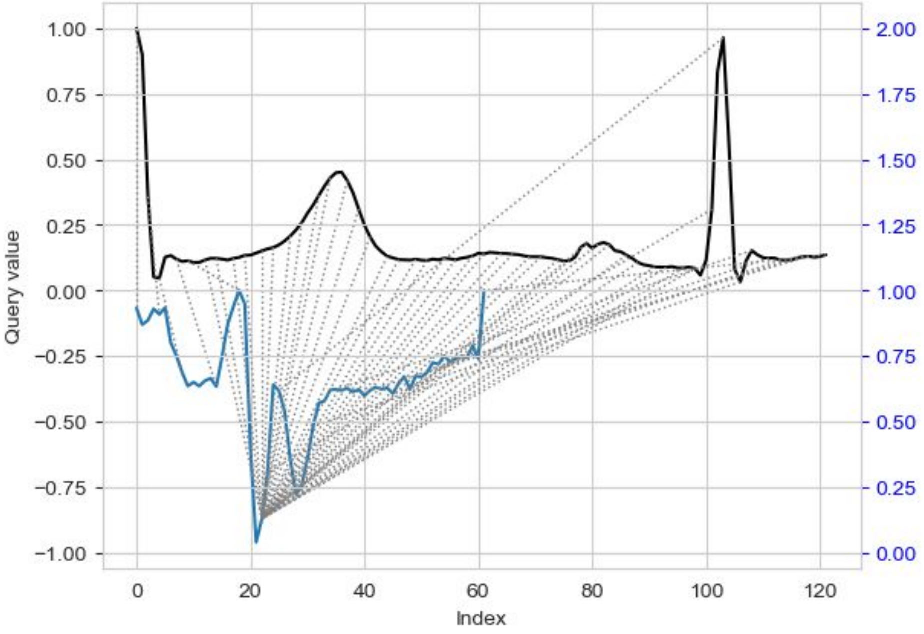
MIMOTO

Example - Normal EKG with DTW



MIMOTO

Example - Abnormal EKG with DTW



MIMOTO

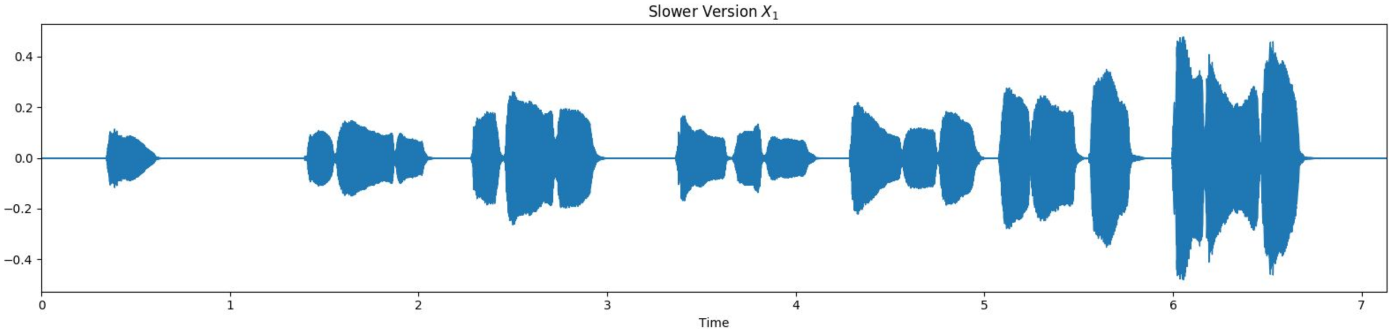
Stock Market - Tesla Momentum Pause Patterns

https://medium.datadriveninvestor.com/creating-a-momentum-trading-scanner-with-dynamic-time-warping-2a4e7ceb1e1c

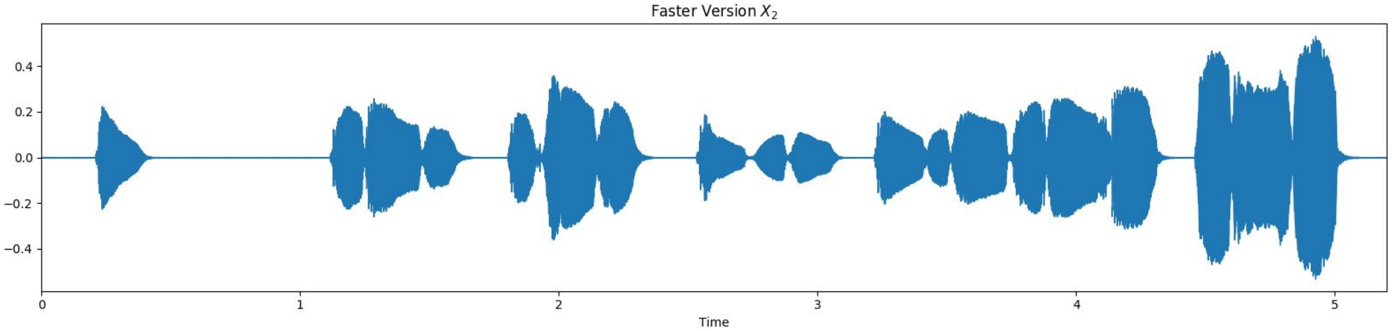
MIMOTO

Stock Market - Tesla Momentum Pause Patterns with DTW

Audio

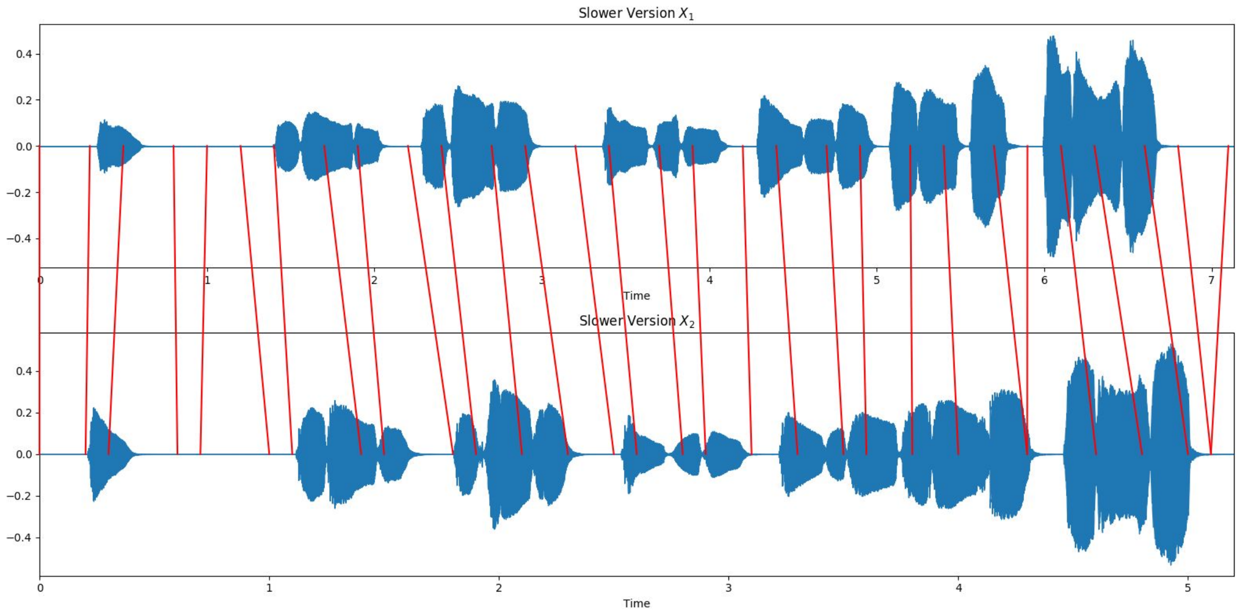


Slower Version X_1
0.4
0.2
0.0
−0.2
−0.4
0
1
2
3
4
5
6
7
Time



Faster Version X_2
0.4
0.2
0.0
−0.2
−0.4
0
1
2
3
4
5
Time

Audio with DTW

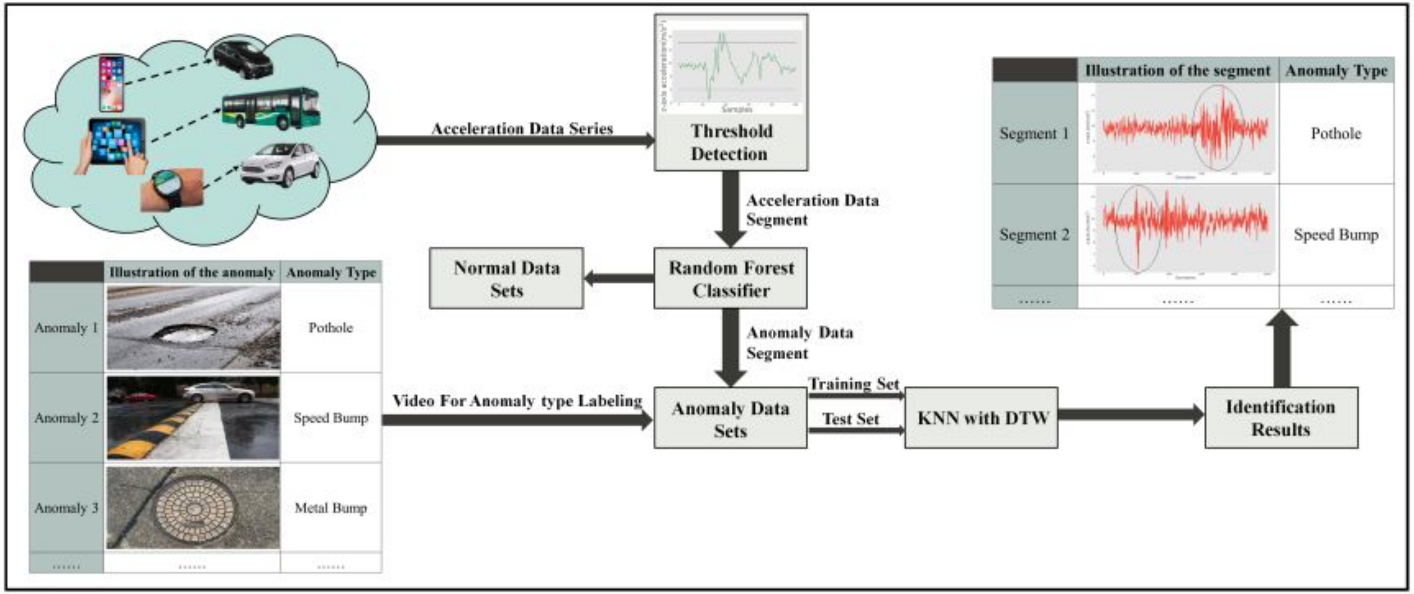


https://librosa.org/librosa_gallery/auto_examples/plot_music_sync.html

MIMOTO

Machine Learning Applications

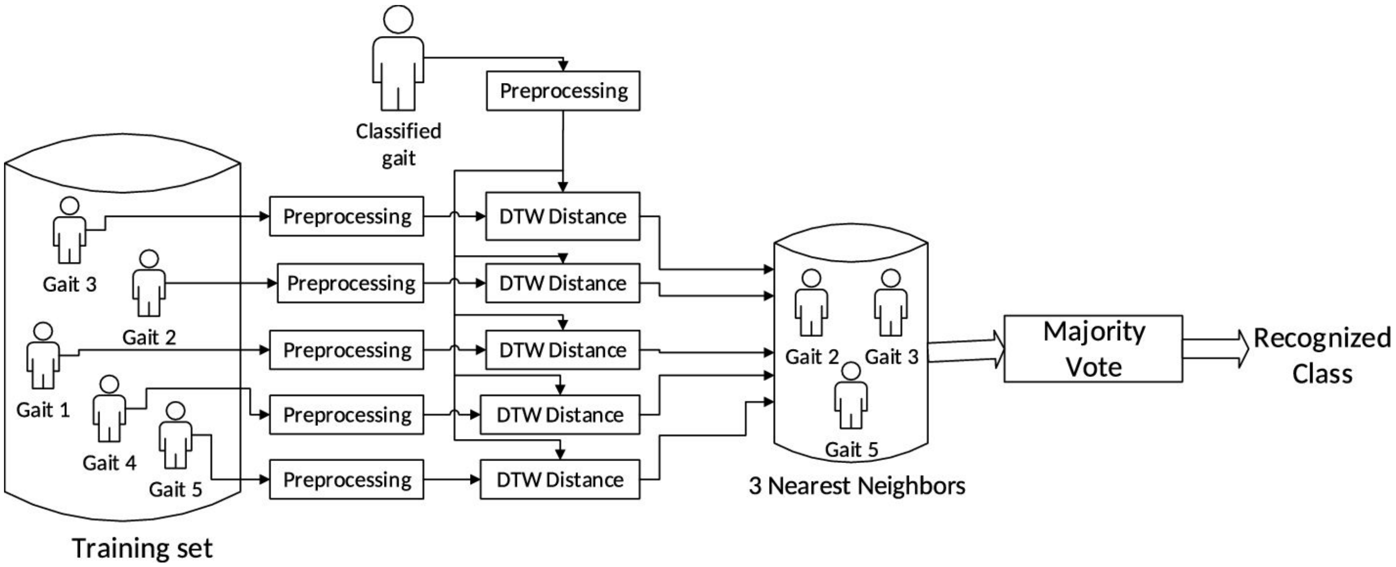
Road Anomaly Detection



Z. Zheng et al., "A Fused Method of Machine Learning and Dynamic Time Warping for Road Anomalies Detection," in IEEE Transactions on Intelligent Transportation Systems, vol. 23, no. 2, pp. 827-839, Feb. 2022, doi: 10.1109/TITS.2020.3016288.

Motion Capture Data

[Switonski, A., Josinski, H. & Wojciechowski, K. Dynamic time warping in classification and selection of motion capture data. *Multidim Syst Sign Process* **30**, 1437–1468 (2019). https://doi.org/10.1007/s11045-018-0611-3](https://doi.org/10.1007/s11045-018-0611-3)



MIMOTO

Assessment of Pulsatile Signals

Li, Qiao, and Gari D. Clifford. "Dynamic time warping and machine learning for signal quality assessment of pulsatile signals." *Physiological measurement* 33.9 (2012): 1491.

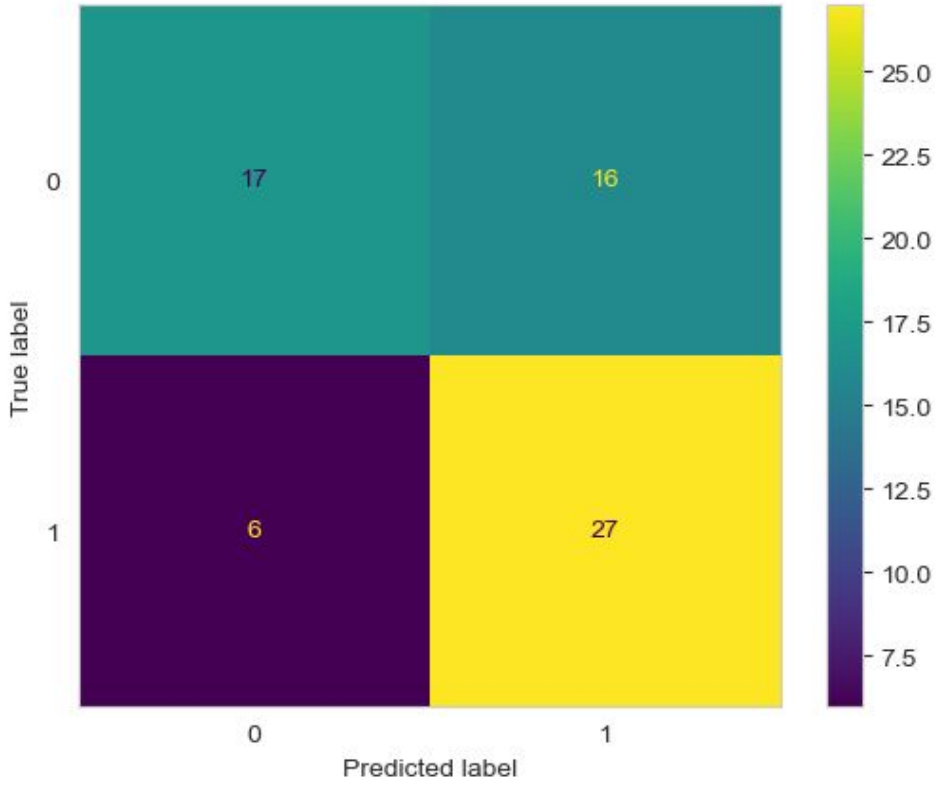
1. Resampled heartbeats to same length
2. Calculated correlation between beats
3. Used correlation as a feature in a neural network to assess the clinical utility of the waveforms

Classification as normal or abnormal EKG using data above

1. One ‘normal’ and one ‘anomaly’ heartbeat per patient
2. Limited dataset to first 100 patients
3. Calculated median dynamic time warping distance between each ‘normal’ and all other ‘normal’ patients
4. Calculated median dynamic time warping distance between each ‘abnormal’ and all other ‘normal’ patients
5. Optimized k in k nearest neighbors using recall

Results

```
Accuracy: 0.667
Precision: 0.628
Recall: 0.818
F1: 0.711
```



MIMOTO

Take-Aways

1. Understand what dynamic time warping is
2. General understanding of the available packages (python and R)
3. Examples of how it boosts outcomes when combined with machine learning techniques

REPO: https://github.com/kimkraunz/odsc_west_2023

MIMOTO

Questions?