Dynamic Time Warping For Brain Meg Analysis

Sharmistha[1] Jat, Tara Pirnia[2], Tom Mitchell[3]

[1]Indian Institute of science, Bengaluru India

[2]Center for Neural Basis of Cognition, Carnegie Mellon University

[3]Machine Learning Department, Carnegie Mellon University

INTRODUCTION

- Machine learning models are used to predict brain activity from stimulus features for discovering how neural signals encode information[1][2]. Testing of these models is critically dependent on the choice of metric to compare predicted time series and the observed data.
- Using Euclidean distance to measure how well the predicted activity matches the observed data is unrealistic given the natural variance in timing across a single subject's response to the same stimulus.
- Dynamic time warping (DTW) method does not assume perfect alignment between the predicted and observed activity and can adjust for small shifts and stretches in time between two time series [3].
- Our work uses DTW method to compare two time series, neural activity recorded using magnetoencephalography (MEG) and predicted activity from our trained models. We evaluate the performance of our DTW approach with the standard Euclidean metric.

METHODS

Datasets: Two data sets, D1 (n= 4 subjects) and D2 (n = 9) [4].

Experiment: Noun stimuli decoding from vector (GloVe and Word2vec) representations using linear ridge regression [5]. Distance were calculated after z-scoring both the predicted and observed time series data.

DTW distance: Compute the DTW distance with constrained window to a fixed +/-5 millisecond in time series.

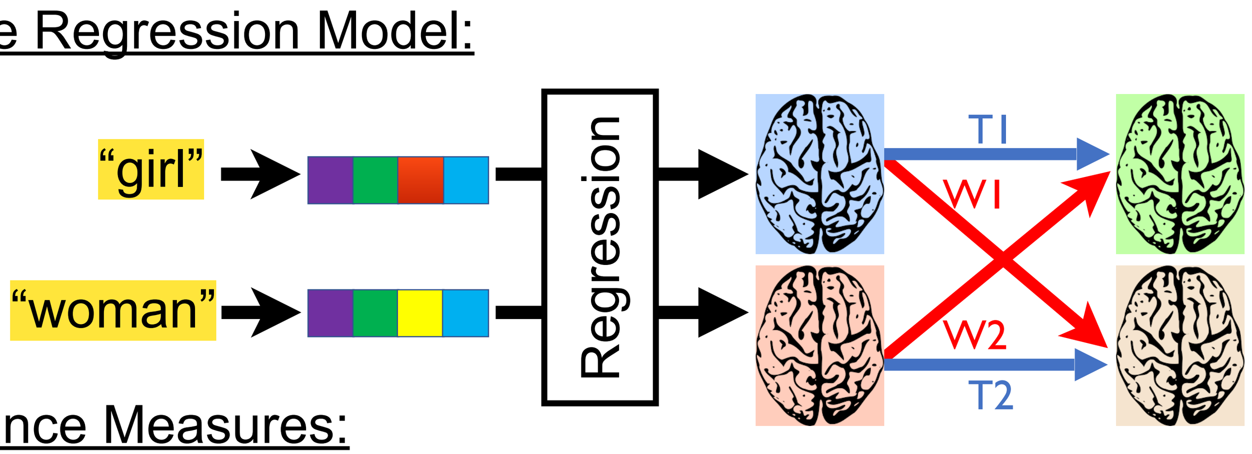
Accuracy: 5 fold cross validation was used.

Sample Stimuli:

"the woman was encouraged by the boy"

"the girl encouraged the woman"

Ridge Regression Model:



Distance Measures:

DTW: $d(w_k) = distance(x_i, y_j)$

$$d_{DTW}(x,y) = min_W\left(\sum_{i=1}^{K=500} d(w_k), W =< w_1, \ldots, w_K >\right)$$

Alignment of (x_i, y_j) is $w_k = (x_i, y_j)$ on path W

Euclidean:

$$d_{eucl}(x,y) = \sqrt{\sum_{i=1}^{M=306}\sum_{j=1}^{L=500}(x_{ij}-y_{ij})^2}$$

Accuracy Measure:

$T1 + T2 < W1 + W2 \Rightarrow \text{Accuracy} = 1$

$T1 + T2 > W1 + W2 \Rightarrow \text{Accuracy} = 0$

$T1 + T2 = W1 + W2 \Rightarrow \text{Accuracy} = 0.5$

Figure 1. Method Schematic for the experiment to decode noun stimulus in brain activity using GloVe embedding features for a word. The MEG signal is preprocessed using temporal signal space separation (tSSS), low-pass filtered to 150Hz with notch filters at 60 and 120Hz, and down-sampled to 500Hz. (A) Sample sentences from the experiment. (B) Regression from word features to predicted brain activity. (C) Distance measure formulae for input x,y. (D) Accuracy computation for a word pair (x1,y1), T1 = distance(x1 , y1), W1 = distance(x2, y1).

ACKNOWLEDGEMENTS

This work is supported by generous grants from Government of India scholarship, BrainHub scholarship and travel grant from oHBM conference to Ms. Sharmistha Jat. Additional funding to attend this conference was provided by the CMU GSA/Provost Conference Funding.

RESULTS

- Our results in the figure below show that the decoding accuracy for both GloVe and Word2vec stimulus representations improves when DTW distance is used as compared to Euclidean distance.
- This improvement in accuracy is encouraging in probing the time series shift within a single subject response.

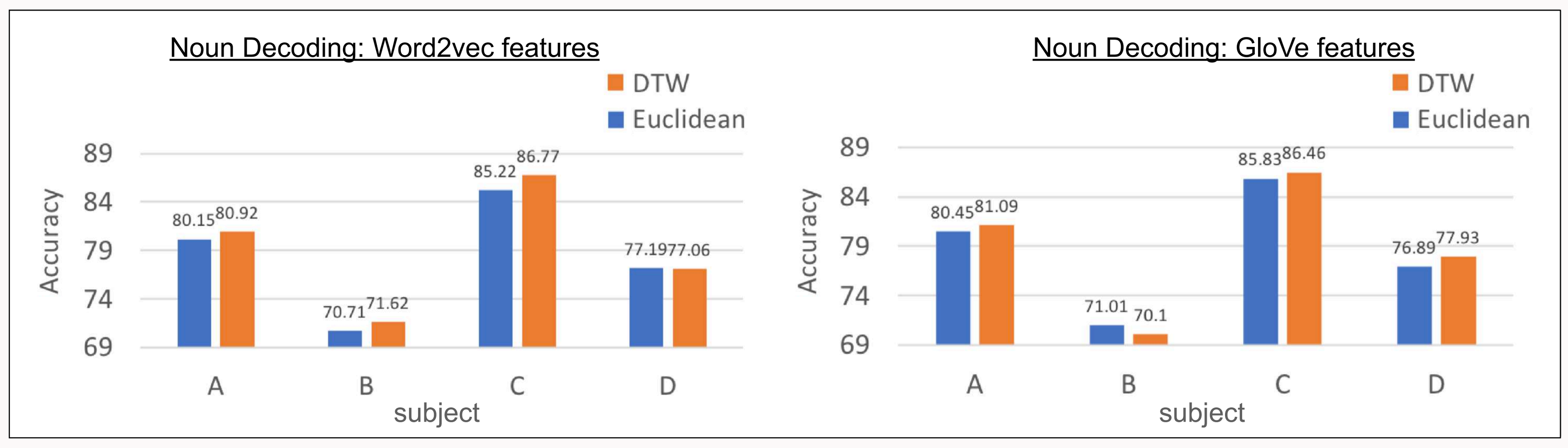


Figure 2. Plot shows accuracy of noun decoding experiment using multiple distance measure, we use GloVe [6] and Word2Vec features to decode the stimulus, from the plot we observe that the DTW warping distance gives significant decoding accuracy.

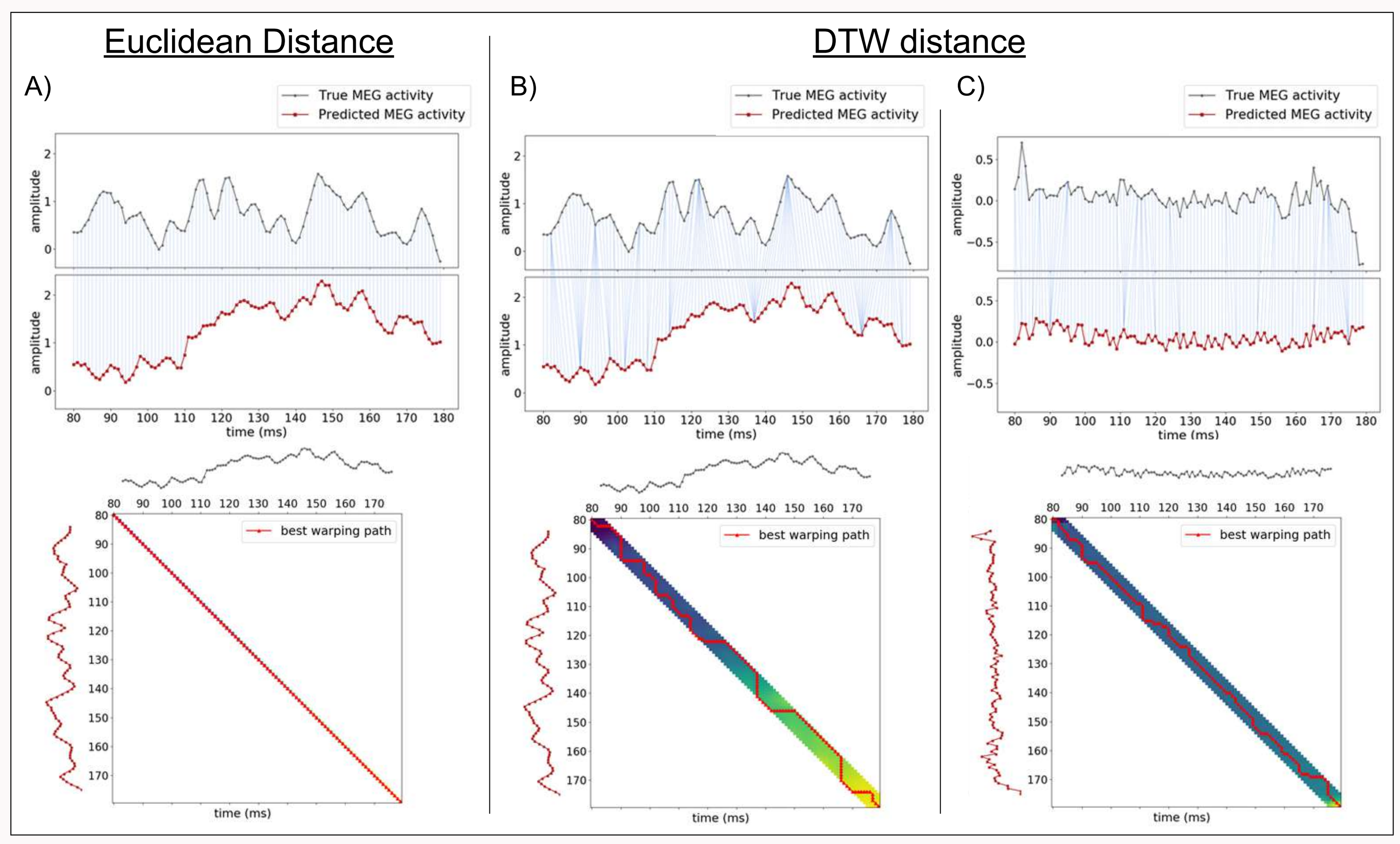


Figure 3. Mapping of points between predicted and true brain activity over time for Euclidean and DTW distance measures. A single sensor for the noun "boy", sampled for 100ms. (A) & (B) show the sensor data from subject 'A', while (C) shows the data from subject 'B'. Bottom row depicts the path between time points used to calculate each distance measure (red line).

DISCUSSION

- Our analysis shows that DTW very significantly outperforms Euclidean distance in classification tasks that require comparing predicted to observed MEG neural activity.
- DTW stretches and shifts data from the sensors and constraining the time warping path.
- Based on this analysis we recommend consideration of using DTW over Euclidean distance in a wide range of classification and prediction problems involving neuroimaging.

REFERENCES

1. Mitchell, Tom M. and Shinkareva, Svetlana V. and Carlson, Andrew and Chang, Kai-Min and Malave, Vicente L. and Mason, Robert A. and Just, Marcel Adam (2008). Predicting Human Brain Activity Associated with the Meanings of Nouns. Science, vol 320, 1191-1195.
2. Wehbe, L., Vaswani, A., Knight, K., & Mitchell, T.M. (2014), 'Aligning context-based statistical models of language with brain activity during reading', EMNLP, vol. 14, pp. 233-243.
3. Berndt, D.J., Clifford, J. (1994), 'Using dynamic time warping to find patterns in time series', Proceedings of the 3rd International Conference on Knowledge Discovery and Data Mining (AAAIWS'94), AAAI Press, pp. 359-370.
4. Nicole Rafidi (2014), CMU project report.(https://www.ml.cmu.edu/research/dap-papers/DAP_Rafidi.pdf)
5. Golub, Gene H. and Heath, Michael and Wahba, Grace (1979). Generalized cross validation as a method for choosing a good ridge parameter. Technometrics, vol. 21, 215–223
6. Pennington, J., Socher, R., Manning, C. D. (2014), 'Glove: Global Vectors for Word Representation', EMNLP, vol. 14, pp. 1532-1543.