

# **CS 674 – Data Mining on Multimedia Data**

## **HW 2**

**Due 3/15/17 at 11:59pm**

We discussed several algorithms to find repeated patterns (motifs) in time series data. In this assignment, you will perform motif-based classification on the same time series datasets that you used in HW1, and compare with your DTW-NN-based classifier. You don't need to implement the motif discovery algorithm, though you can if you feel like it. You can either use publicly available code, or modify the code for the algorithm of your choice. What you will need to implement, and to think about, is how to take the motifs discovered, and use them for classification. Some questions to think about include: How do you turn motifs into features? Should you use an algorithm that finds pair motif or range motif (though you could easily find range motifs from pair motifs)? What makes a motif discriminative for classification? How do you select the most discriminative motifs? How many motifs? What classifier to use? If you choose an approximate motif discovery algorithm, some post-processing might be required.

To submit: source code, README, classification accuracy for your motif-based classifier vs. the accuracy from the last assignment, and a report (graphs showing the results, analysis, description of what you did, answers to the questions above, etc.)