Instructions

We ask that you answer a few questions about reading assignment. By design, some of these questions are open-ended and the purpose is to start the thinking process. Some of those questions may already have been answered in recent literature (or in the given paper itself). We encourage you to, both, come up with your own ideas and survey recent papers before answering these questions. A few key points to take care of:

- Answer **two out of the three questions** (you are welcome to answer all of them), unless stated otherwise.

- A lot of questions ask you to suggest changes or make alterations. It would be nice to support it with logical / mathematical arguments. Cite all sources used in the process of coming up with your answer. Figures and equations, if they support your arguments will be appreciated.

Questions

Deep Canonical Correlation Analysis

1. Canonical Correlation Analysis finds correlated representations among two views of data. A shortcoming of this approach is complementary information between the two views are not directly accounted for. In what situations would this be an issue? In other words, would you rather have two representations that are correlated with each other or one joint representation.

2. Authors cover KCCA in the same paper. Would there be a situation where you would believe KCCA works better than DCCA? Also, the authors introduced a new nonlinearity as a replacement for tanh. What are the potential merits of using this activation function over tanh or Sigmoid?

3. Authors suggest that Stochastic Gradient Descent approaches over minibatch are not applicable to DCCA, however Gradient Descent approaches are still possible over the entire training set. Why is that the case? What would happen in your opinion if SGD is in fact used?