

Home Work 3, due Sunday, Nov. 24, 11:59 pm

Another way to encode a corpus is to run a topic allocation model and use the topic distribution for each document as the encoding vector.

Run the LDA algorithm on the reviews with different numbers of topics and use the array of topic distributions as an encoding.

Write your own classification neural net and use the encoding from the topic distributions to train and evaluate the network.

Try different numbers of topics and evaluate the performance for each on the test set

