
CS688: Graphical Models - Spring 2020

Assignment 3 - CRF Training

Assigned: Wednesday, Mar 25th. Due: Friday, Apr 3th, 5:00pm

General Instructions: Please upload **two items** to Gradescope (<https://www.gradescope.com/courses/86501>): (1) a report with your answers (*.pdf*), and (2) a zip file with your code (*.zip*). To help you get started, the full L^AT_EX source of the assignment is included with the assignment materials. For your assignment to be considered “on time”, you must upload a zip file containing all of your code to *Gradescope* by the due time. Make sure the code is sufficiently well documented that it’s easy to tell what it’s doing. You may use any programming language you like. For this assignment, you may not use existing code libraries for inference and learning with CRFs or MRFs. If you think you’ve found a bug with the data or an error in any of the assignment materials, please post a question to the discussion forum. Make sure to list in your report any outside references you consulted (books, articles, web pages, etc.) and any students you collaborated with.

When you submit reports through *Gradescope*, you are supposed to mark what part of the *.pdf* corresponds to each question. Please note that you will lose credit on this assignment if you fail to do this.

Question 1. (100 points) Use **average log conditional likelihood** as the objective function. Implement the objective function, and also its gradient functions (the computation of the partial derivatives of the objective function with respect to each parameter), as derived in Assignment 2. Use your implementation of the sum-product message passing algorithm from Assignment 2 as a subroutine to make your objective and gradient function implementations computationally tractable. Implement the learning algorithm for CRFs by using the numerical optimizer you selected in Assignment 2 to maximize the log conditional likelihood function. Use the first 50, 100, 150, 200, 250, 300, 350 and 400 training data cases to train eight separate CRF models. Answer the following questions.

1.1 (20 points): *Record the total training time in seconds for each of the above training data set sizes. Report your results as a line graph of time in seconds versus training set size. Make sure to label the axes of your plot.*

1.2 (40 points): *Evaluate the prediction error of each model on the complete test set. As in Assignment 2, predict the character with the highest marginal probability for each position of each test word. Report the error rate averaged over all predicted characters in all test words for each model. Summarize your test error results in a line graph showing prediction error versus training set size. Make sure to label the axes of your plot.*

1.3 (40 points): *Evaluate the average conditional log likelihood of the complete test set under each model (for each model, this will be an average of the per-word conditional log likelihoods for each word in the test set). Summarize your results in a line graph showing average conditional log likelihood versus training set size. Make sure to label the axes of your plot.*