Part 1:

This part of HW2 is a group assignment. Look through the Task A and Task B descriptions, and then discuss & submit answers for the following questions with your group. Only one group member needs to submit answers but you should work collaboratively (e.g., maybe draft answers together in a google doc).

1. Come up with a labeling scheme. How many labels do you want to have? Are they binary, or multi-class? It might help to think about / decide how you want to use the resulting model. Write a few sentences on how you came to your decision.

2. You have a budget to hire people to annotate the comments according to your labeling scheme. Write a set of instructions to give to the annotators.
    The format of your instructions is flexible, but as a guide you can consider the following parts:
    a. Brief task description including what they will see & be asked to do (~1 paragraph)
    b. Step-by-step instructions for how to label each example
        • Tips for reading the examples (~1 paragraph): For example, are there specific factors they should consider/look out for? Should they go with their first reaction or look up additional information?
        • Your labeling scheme (~1-2 paragraphs or several bullet points): Make sure you describe your labeling scheme precisely—you don't want people to be labeling examples with different terms, since that will make it hard to reconcile the labels from different annotators. Define terms that are not clear, and maybe include guidelines for what people should do if they're not sure (Should they go with their best guess? Or are you going to include a "Not Sure" label in your scheme?).
    c. Sample labeled examples according to your labeling scheme (e.g. use the examples in your task description doc)

3. Write a few sentences about how you might integrate the annotations into the dataset. Will you average labels of multiple annotators for each example or use some other aggregation? What do you do with examples where there's disagreement among annotators? It's okay if you're not sure, just try to write about what the tradeoffs might be.

4. Are there concerns you have about using technology to address this problem?