ASSIGNMENT C4
CORE METHODS IN EDUCATIONAL DATA MINING
PROFESSOR RYAN BAKER
PROMPT ENGINEERING
YOUR ASSIGNMENT IS DUE DECEMBER 15, 1159PM USA EASTERN
YOUR RESPONSE POSTS ARE DUE DECEMBER 18, 1159PM USA EASTERN

The goal of this assignment is to improve a prediction model using prompt engineering. The existing code to build the prediction model (all variants) is found at https://github.com/JZ2655/CSCL23/blob/main/CSCL23_feedback%20detectors.ipynb

This model is described in

Zhang, J., Baker, R.S., Andres, J.M.A.L., Hutt, S., Sethuraman, S. (in press) Automated Multi-Dimensional Analysis of Peer Feedback in Middle School Mathematics. Proceedings of the International Conference on Computer Supported Collaborative Learning.

Which can be found at https://learninganalytics.upenn.edu/ryanbaker/ISLS23_annotation%20detector_short_submit.pdf

This paper was submitted right as GPT-3.5 was becoming available, so it did not use the current generation of large language models (it used the Universal Sentence Encoder, which is great).

Your goal is to use prompt engineering (or fine-tuning, or whatever you want) on a contemporary large language model of your choice (GPT, Claude, LLaMA, Bard, or something not yet available when I wrote this assignment at the beginning of the semester) to improve this model of **Commenting on the process (CP)**. You can use the LLM to improve the input to the model training algorithm. If you do this, the model training algorithm should be a neural network set up in the exact same fashion as Zhang et al. Or you can just use the LLM to obtain inferences. (In other words, do not try other algorithms like XGBoost, because that's not the goal of the assignment).

Your goal is to do better than Zhang et al. on this construct, using the same cross-validation scheme and random seed as that paper. However, doing better than Zhang et al. is not a requirement for a good grade. The requirement is to use a large language model in an appropriate fashion, and to try.

Please post to the forum, in a new thread within the CA4 folder:

- Text explaining how you completed the assignment
- Evidence of model goodness, when the model is applied to new students (use AUC ROC and the same cross-validation scheme and random seed as Zhang et al.)
- Along with files
    - The data set you input into the neural network, if different than the original data set (and if you used a neural network)
    - Any prompts you used to generate that data set or to generate outputs
    - The ipython notebook(s) (or other code) that you used at any point

Solutions will be graded on completeness and comprehensibility, whether you correctly and validly apply the method you choose to this data, and whether the methods you chose fit the requirements of this assignment.

BONUS: The student who succeeds in producing the detector with the best AUC ROC, under correct cross-validation, gets the bonus.

PART TWO: YOUR RESPONSE POSTS

After completing your own assignment, you are expected to also provide substantive comments on at least four other students' submissions, as a response within that student's assignment thread. For these posts, there is no length requirement, but the posts must offer a critical and meaningful perspective on how that student did the assignment. (i.e. "Great job! You did really awesome!" and "Terrible! You totally messed up!" are insufficient)

This is not just for the benefit of the student whose solution you are commenting on. Seeing how other students did this assignment will be informative to you as well.

Although there is no requirement to do this, you are encouraged to give feedback to students who have received fewer feedback responses so far – i.e. I would like to avoid having one student get feedback from every classmate, and another student get feedback from no one. Thanks.