Question: Language Models

a. Consider a simple fixed-window neural language model (covered in class) that is being trained on a large corpus of English text.
   i. Describe how one epoch of training looks like (e.g., via a list of steps or a pseudo-code).

   ii. What are the main advantages of using an RNN-based language model over the fixed-window neural model?

b. Consider a system that classifies short news articles into one of several categories (sports, science, medical, ….). The system feeds the short article word-by-word into an RNN, and then computes the predicted class based on the last hidden state of the RNN.
   i. You observed that the model has mistakenly classified the following short article as "medical":
   Jane Doe won the Formula One championship, denying John Doe a record eighth, with a last-lap overtake to win a season-ending Abu Dhabi Grand Prix on Sunday despite a recent two-day hospitalization.
   Explain how your choice of model architecture could be leading to the misclassification.

   ii. How would you modify your current RNN-based architecture to improve the performance of your model in cases such as the one described in (i)? Justify your answer.