Name:

NetID:

Note: This quiz will be a set of practice problems for the upcoming exam. Complete this at your leisure outside of class but please bring in an **INDIVIDUAL** copy of the quiz **ON PAPER** to turn in at the start of the exam. While doing the work feel free to work with your classmates, but I'd like you all to turn in a copy individually. You don't necessarily need to print and fill out this exact document, just writing the answers on your own piece of paper will be fine.

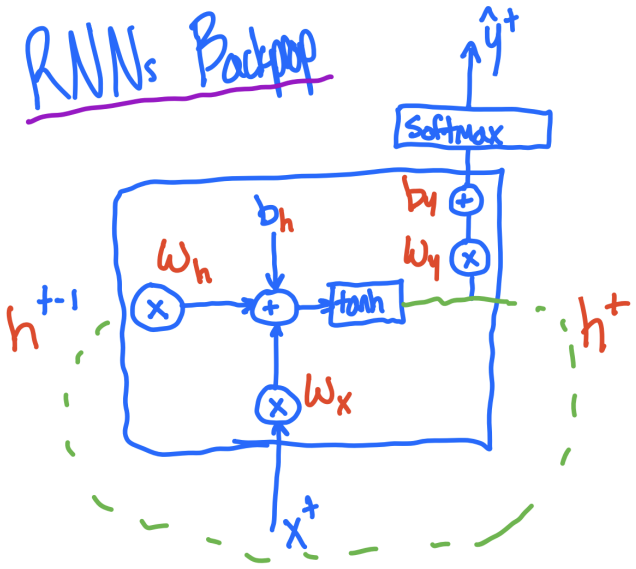
1.) What is the bias/variance tradeoff?

2.) How does k-Nearest Neighbors classify points when used for classification?

3.) What is the purpose of the Gaussian distribution in a Naive Bayes classifier?

4.) What is polynomial feature expansion?

5.) Given the network diagram, draw the computation graph for the network for the sentence “Good Luck” with a word level tokenizer, only producing output for the final token. (5 pts.)



6.) What are tokens and embeddings in relation to NLP tasks?

7.) Convolutional Networks use filters or kernels to process images. What do these look like and what purpose do they serve?

8.) What are the two types of pooling layers? Why do we use them?

9.) What is the difference between Djikstra’s and A*?

10.) We say Attention(Q, K, V) is a “dictionary lookup”, explain this metaphor.

11.) What is layer normalization and why do we use it in a transformer?

12.) How are minimax and alpha/beta pruning related? Do we still have to use minimax if we use alpha/beta pruning?

13.) What are the differences between gradient descent, stochastic gradient descent, and mini-batch stochastic gradient descent?

14.) What is a linear transformation and how does it relate to deep learning?

Bonus.) The draft website for next semester is available at nextai.williamtheisen.com I'd love it if you were willing to look at it and give me any thoughts you may have. Our TA, Tom, suggested that if you turned in the exam practice packet you could get points on the exam which I thought was a really good idea!