

2.

- a) The five words with the highest values are ['nations', 'free', 'war', 'government', 'world'] with values [0.00318534, 0.00330975, 0.00371982, 0.00394117, 0.00470717] respectively. The five words with the lowest values are ['jobs', 'years', 'work', 'care', 'americans'] with values [-0.00525501, -0.00383007, -0.00357746, -0.00342131, -0.00328322] respectively.
- b) The confidence of the classifier when it is both correct and incorrect are always close to 100%. The lowest the confidence dropped when it was incorrect was ~91%, which is still very confident. What can contribute to the a high confidence in the model is that it only looks at the data that it is given so when the model predicts a set with new data it uses the new data incorrectly. However, the model estimates what it does is correct based of the past data. Additionally, the model treats all data independently so if data is not independent, it can not distinguish the data dependencies and incorrectly predicts the outcome, but thinks that it is correct.
- c) The problem with having a high confidence when the value is incorrect is that the algorithm is giving false information to the user. In the case of the diagnosis algorithm, the algorithm confidently diagnosed the patient incorrectly which leads to the patient getting a treatment that does not help or can even worsen the patient. Furthermore, it could lead to doubt in the algorithm so that the next time the algorithm diagnosis a patient and it is correct, the physician might disregard the result. A benefit with a confident classifier with a correct prediction could be that it can quickly give a diagnosis and save the patient time and money or even catch and treat a disease early before it gets bad.
- d) The regulator can help prevent overfitting of the function. This can be done by limiting the confidence levels and adding different conditions while training. This reduces the likelihood and add weights to help prevent overfitting of the function.