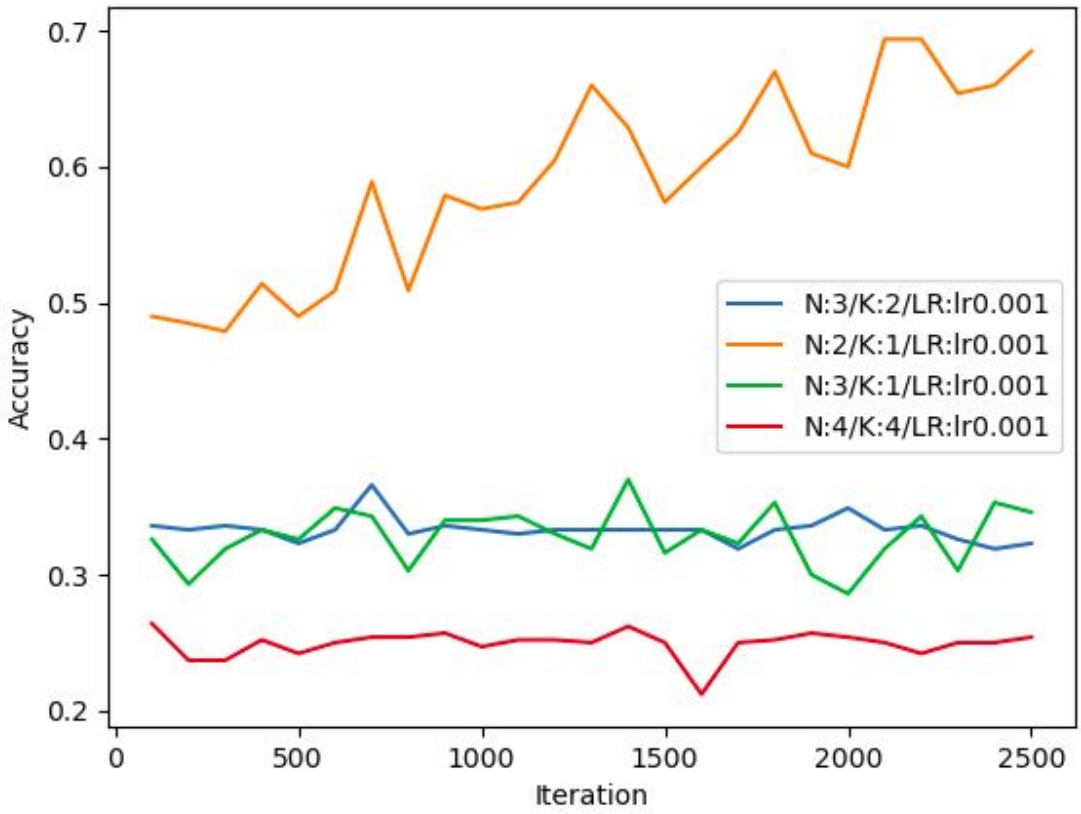
Emre Girgin - 2016400099

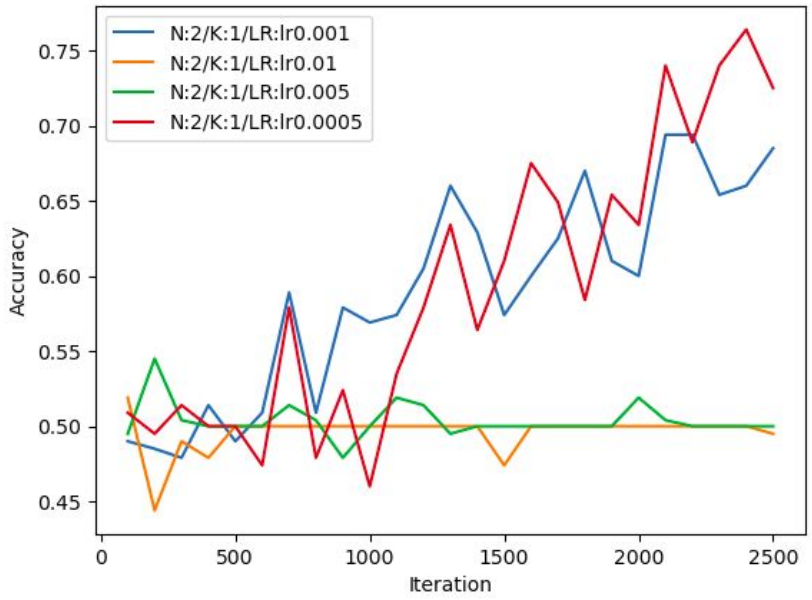
Question 3: (see question3.sh)



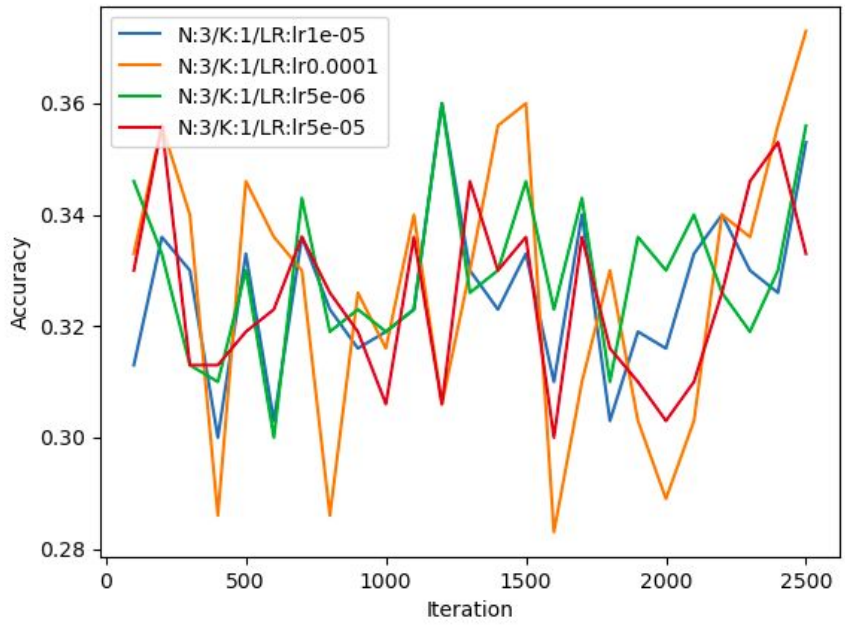
Observation:

- It seems only the N: 2 K: 1 model seemed to learn in 2500 iterations.
- The other configurations did not learn anything, yet.
- The remaining models seem predicting randomly. (Accuracy = 1/N)
- However, since the first configuration started to learning, it proves the code is working.
- The other configurations may need more iterations to learn but I could not complete them due to hardware limitations.

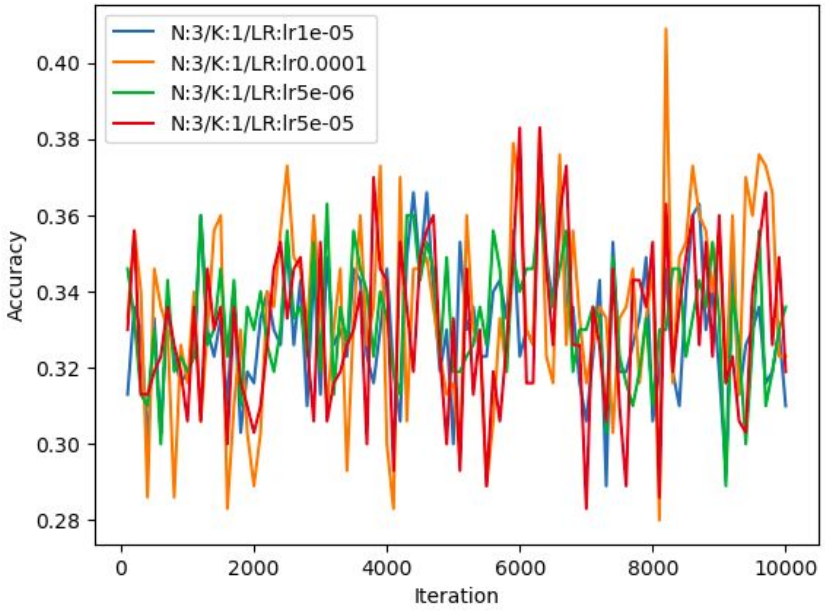
Question 4.a: (see question4.sh)



I first tried different learning rates for the configuration of N: 2 K:1 since it is the only learning configuration I had. Then observed that after the learning exceeds 0.001 the learning got worse.



Then I tried the N:3 K:1 configuration with lower learning rates but could not obtain a pattern in 2500 iteration.

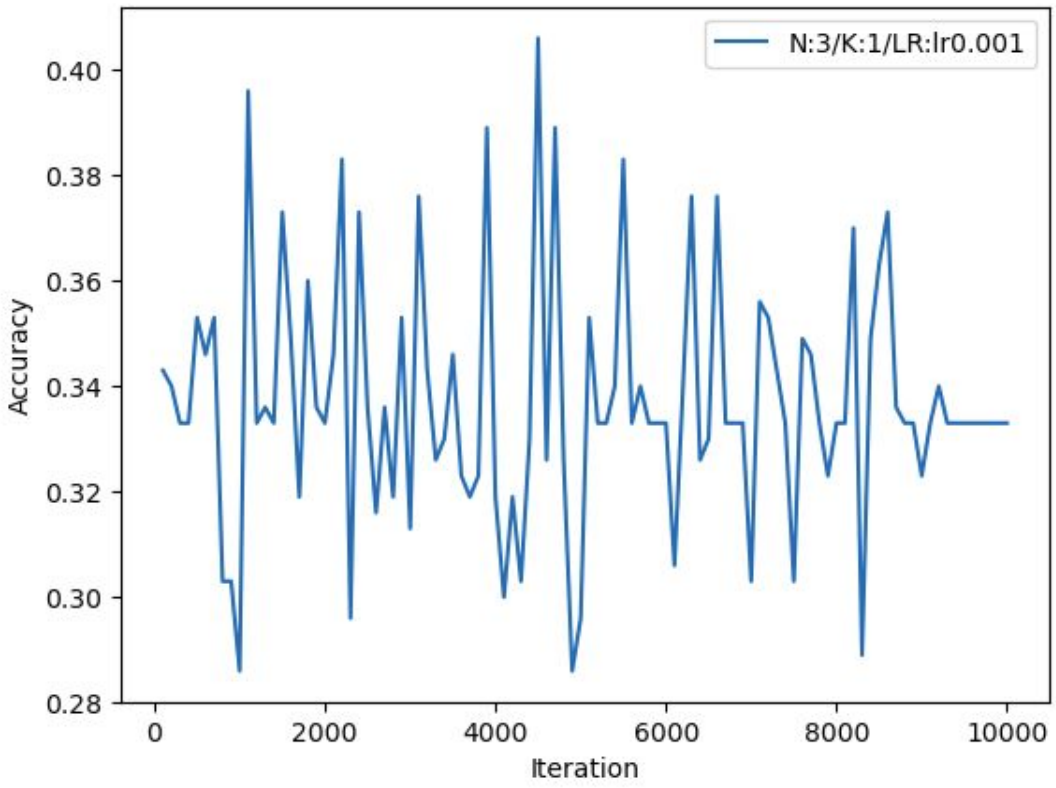


I trained the models for the 10000 iterations but there was not much improvement.

Question 4.b: (2 layer 2D convolution) (see question4_b.sh)

I have added to conv layers before the LSTMs:

- Layer 1 :
    - # filters : 11
    - Kernel size : 3x3
    - Padding : Same
    - Activation : Relu
- Layer 2 :
    - # filters : 1
    - Kernel size : 3x3
    - Padding : Same
    - Activation : Relu



Unfortunately, the convolution layers do not seem to help much.