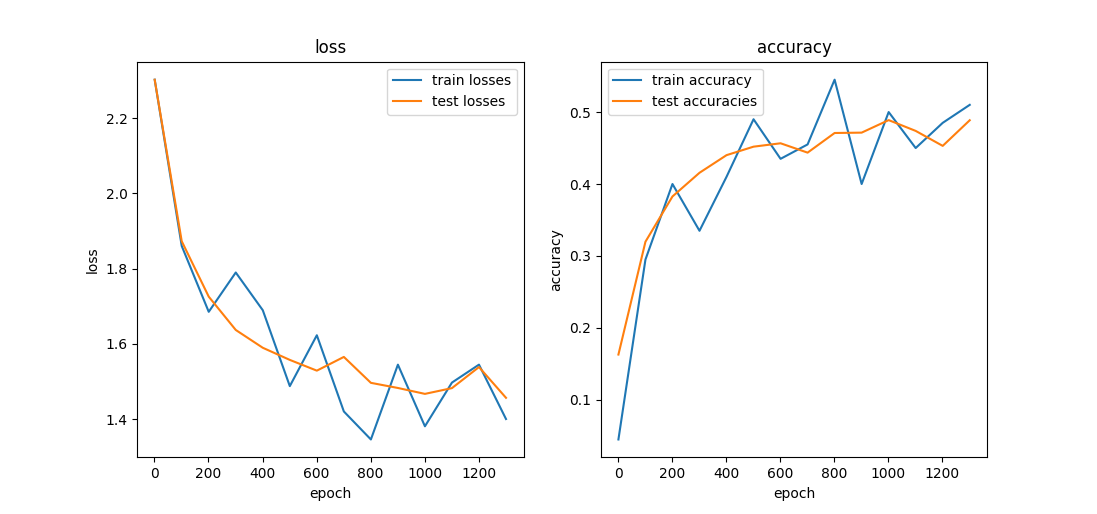
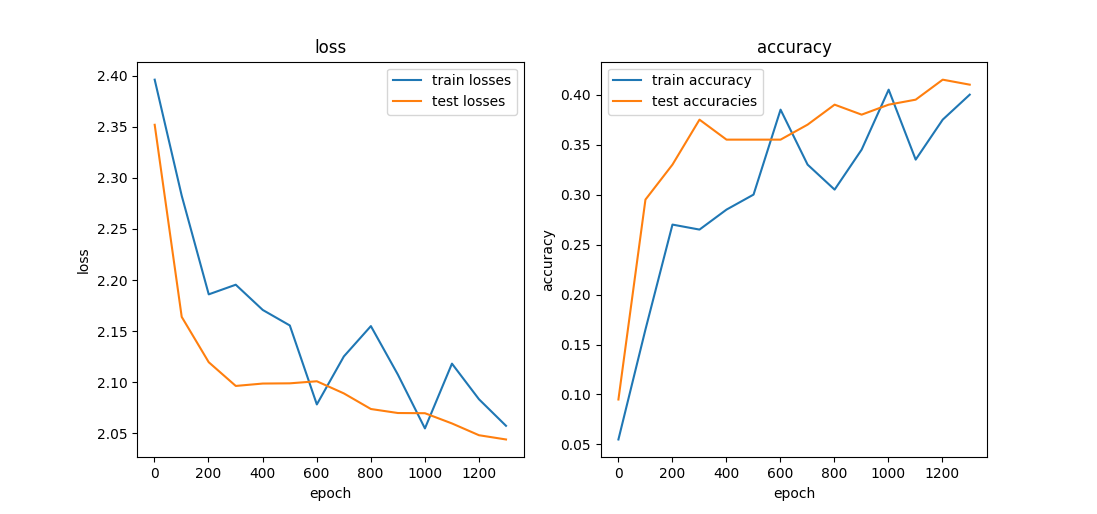
1.2 numpy



2 pytorch



goal is to get the best test accuracy you can. You should be able to get at least 0.52 accuracy on  
the test set but we challenge you to improve this.

|  |  |
| --- | --- |
|  |  |
|  |  |

Briefly discuss about the benefits and/or the drawbacks of using the function *Tanh,*

instead of the activation function ELU you used before.

Benefit of Tanh: it has a strong gradient. Strong gradient is necessary for optimizing parameters. Output are centered around 0.

Drawback of Tanh:

Vanishing gradient—for very high or very low values of X, there is almost no change to the prediction, causing a vanishing gradient problem. This can result in the network refusing to learn further, or being too slow to reach an accurate prediction.

Activation Func 优缺点比较

<https://www.kaggle.com/getting-started/157710>

<https://prateekvishnu.medium.com/activation-functions-in-neural-networks-bf5c542d5fec>