hw3_report

Describe what problems you encountered and how did you solve them when implementing the basic and advanced functions.

1. Formula Interpretation: use open AI to convert formula into Python code
2. Got confused abt the cache value: print out the state of it
3. assertion error of the shape of dZ and Z: incorrect formula interpretation
4. Overflow on backward sigmoid function: set the boundary of exp, i.e.

   replace:

   ```
   s = 1 / (1 + np.exp(-Z)
   ```

   with:

   ```
   s = 1 / (1 + np.exp(-np.maximum(-Z, -700)))
   ```

Briefly describe the structure of your binary and multi-class classifiers.

- Binary: input layer (w/ num. of features) → hidden layer (relu, 64 neurons) → hidden layer (relu, 32 neurons) → output layer (softmax, 2 classes)
- Multi-class: input layer (w/ 28 * 28 neurons) → hidden layer (relu, 128 neurons) → hidden layer (relu, 64 neurons) → output layer (softmax, 100 classes)

Describe effort you put to improve your model (e.g., hyperparameter fine tuning, losses' impact on the result).

- I found that the loss is asymptotical to 203.4, the ratio of learning rate and num. of iterations in basic part is 0.0001/3300
- The ratio in advanced part is 0.01/500 (default value)