

Homework 1

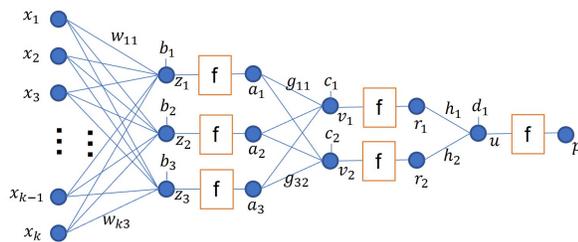
DS36061 Artificial Intelligence, Fall 2022
 School of BioMedical Convergence Engineering, PNU
 Due: Oct. 10. 23:59 (PM)

I. REMARK

- You must describe your works in answer sheets to get full credit.
- Please submit all code files (Google Colab ipynb files) so that TA can easily run the codes.
- If you disobey the deadline, you will get only 50 % of your full score!!!!
- If you just copy your colleague's, you will get 0 point.
- The perfect score is 100 points. The bonus question is not mandatory!!.
- Answer using Korean or English.

II. PROBLEM SET

- 1) The task is to implement a classifier (logistic regressor) for breast cancer data. The deep-learning model is represented by the figure below. It is a fully connected network having two hidden layers. The function f represents the Sigmoid function.



- a) Load a dataset by importing load_breast_cancer from "sklearn.dataset". Check the number of samples (patients). Also, check the number of features for each sample. [5 points]
- b) Split the dataset into training-set, validation-set and test-set using sklearn. (Ratio: 6:2:2). [5 points]
- c) Pre-process (normalize) the three datasets for scaling the range of each feature. [10 points]
- d) The model must use the cross entropy as a loss function. Also, the parameters of the model must be updated by stochastic gradient descent. Find the derivative of the loss with respect of each parameter. [30 points]

- e) Make a class for implementing the deep learning model using results in d). Note that the model is fed by all features of each sample. Like sections 1-6 in our textbook, you need to make your own class and functions without adopting packages (libraries) such as "sklearn.neural_network". [30 points]
- f) Plot training error and validation error over epoch for checking learning curves. The error indicates the mean loss using all samples. Discuss whether overfitting or underfitting happens.[10 points]
- g) Determine hyper-parameters so that the model provide best performance. After learning, please test your trained model. What is the mean loss using all samples in the test set. What is the score (the number of correct predictions / the number of all samples)? [10 points]

- 2) [Bonus points!!] What if the nonlinear functions in the two hidden layers are rectified linear unit (ReLU) functions? You can easily find the equation or graph of ReLU by just googling. Repeat the process (a-g) by converting the Sigmoid functions to ReLU functions. The function in the output layer must be Sigmoid function.