

Midterm Exam

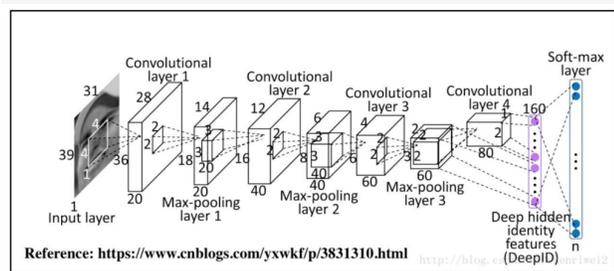
DS36061 Artificial Intelligence, Fall 2022
 School of BioMedical Convergence Engineering, PNU
 Oct. 20. 09:00 - 12:00

I. REMARK

- 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 just copy your colleague's, you will get 0 point.
- Answer using Korean or English.

II. PROBLEM SET

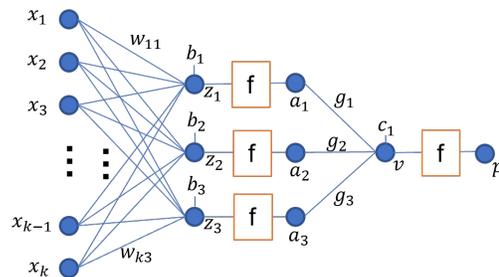
- 1) Answer the following questions [10 points].
 - a) Explain under-fitting and over-fitting in detail.
 - b) Explain L1 and L2 regularizations. Explain why they are helpful for preventing over-fitting?
- 2) Answer the following questions about deep learning [25 points].
 - a) Explain early-stopping during a training process.
 - b) Explain why it needs split a data set into training, validation and test data set.
 - c) Explain stochastic, mini-batch and batch gradient descents.
 - d) Explain why deep learning models use non-linear functions at hidden layers as activation functions.
 - e) Explain why deep learning models for logistic regression (binary classification) has a sigmoid function at an output layer (last layer).
- 3) Answer the following questions concerning CNN [15 points].



- a) Explain the convolution operation. Explain why it has a no problem if CNN adopts the cross-correlation instead of the convolution.
- b) Explain the advantages of the neural network which has convolution layer instead of dense layer.
- c) See the above CNN model. Count the total number of trainable parameters. You can assume that the

model has no bias parameters.

- 4) [Colab] 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 one hidden layer. The function f represents the Sigmoid function [50 points].



- a) Load a dataset by importing `load_breast_cancer` from `"sklearn.datasets"`. Check the number of samples (patients). Also, check the number of features for each sample.
- b) Split the dataset into training-set, validation-set and test-set using `sklearn`. (Ratio: 6:2:2).
- c) Pre-process (normalize) the three datasets for scaling the range of each feature.
- 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.
- 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"`.
- 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.
- 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)?