Fundamentals of Machine Learning (Fall 2022)

Homework #2 (120 Pts, Due date: Oct 5)

Student ID _____

Name _____

Instruction: Once solving the problems, submit two files as follows.

- **'ML_HW2_YourName_STUDENTID.zip'**: Compress the 'models' folder including 'models/SoftmaxClassifier.py' and 'models/LogisticRegression.py.'

- **'ML_HW2_YourName_STUDENTID.pdf'**: Convert your document into a pdf file.

1. Solve the following problems.

(a) **[10 pts]** Suppose we have the following samples from the Bernoulli distribution. Calculate the parameter p using the maximum likelihood estimation method.

$$\{0, 1, 1, 1, 1, 0, 1, 1, 1, 0\}$$

Answer:

(b) [10 pts] Suppose we have the following samples from the Gaussian distribution. Calculate the parameter μ and σ^2 using the maximum likelihood estimation method.

$$\{0.3, 0.5, 0.7, 0.8, 1.2\}$$

Answer:

Instruction: For problems 2 and 3, we provide two classification datasets, the Banknote authentication dataset and the Litmus dataset. The banknote authentication dataset is used for binary classification. It consists of 4 features (e.g., the variance of image, skewness, kurtosis, and entropy) to predict authentication for banknotes. The Litmus dataset is used for multi-class classification, representing a pH scale from 0 to 14. It consists of 3 features (e.g., blue, red, and green). The detailed information for each dataset is as follows.

Dataset	# of training data	# of test data	# of classes	Details
Banknote	1,029	343	2	link
Litmus	488	165	15	-

NOTE 1: You should write your codes **only in 'EDIT HERE.'** Once you complete your implementation, run the checker code (**'0_LogitisticRegression_Checker.py'** or **'1_SoftmaxClassifier_Checker.py'**) to validate if your code is executed correctly.

NOTE 2: You may need to install NumPy and Matplotlib libraries.

NOTE 3: Please carefully read the comments in the code.

2. **[Logistic regression]** Write your code to implement logistic regression. (Default hyperparameter settings for (b), (c): Epoch = 50, Batch_size = 512, learning_rate = 0.1)

(a) **[30 pts]** Implement functions in 'models/LogisticRegression.py'. (**'forward'**, **'compute_grad'**, **'_sigmoid'**, and **'eval'** respectively). Given a mini-batch data (X, Y), the error function for a mini-batch is defined as follows:

$$E(\mathbf{w}) = -\frac{1}{|\mathcal{B}|} \sum_{(x,y) \in \mathcal{B}} y * \log(\hat{y}) + (1 - y) * \log(1 - \hat{y}),$$

$where \; \hat{y} = sigmoid(\mathbf{w}^T \mathbf{x}), |\mathcal{B}| \; \textbf{is the number of the mini} - \textbf{batch samples}.$

**Fill in your code here. You also have to submit your code to i-campus.**

Answer:

(b) **[5 pts]** For the Banknote dataset, draw the plots by adjusting the **learning rate**. (Let the other hyperparameters fix as default hyperparameter settings.) The x-axis is the value of hyperparameters searched, and the y-axis is the accuracy score. Try at least five different values and explain your results.

Answer:
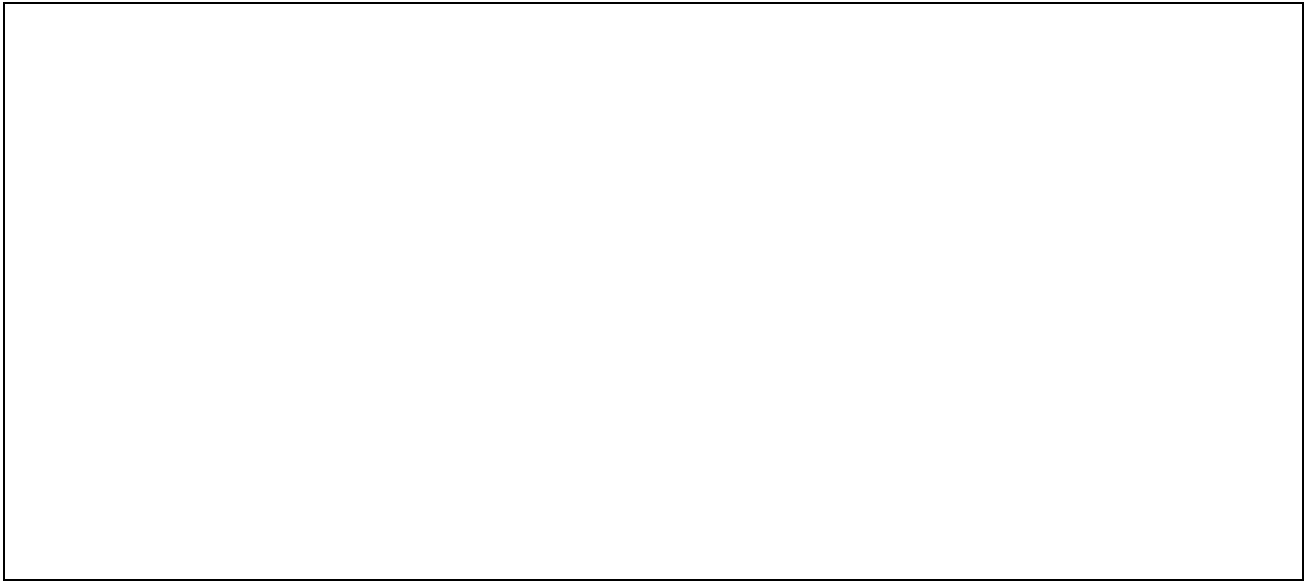
(c) [5 pts] For the Banknote dataset, draw the plots by adjusting the number of **epochs**. (Let the other hyperparameters fix as default hyperparameter settings.) The x-axis is the value of hyperparameters searched, and the y-axis is the accuracy score. Try at least five different values and explain your results.

Answer:
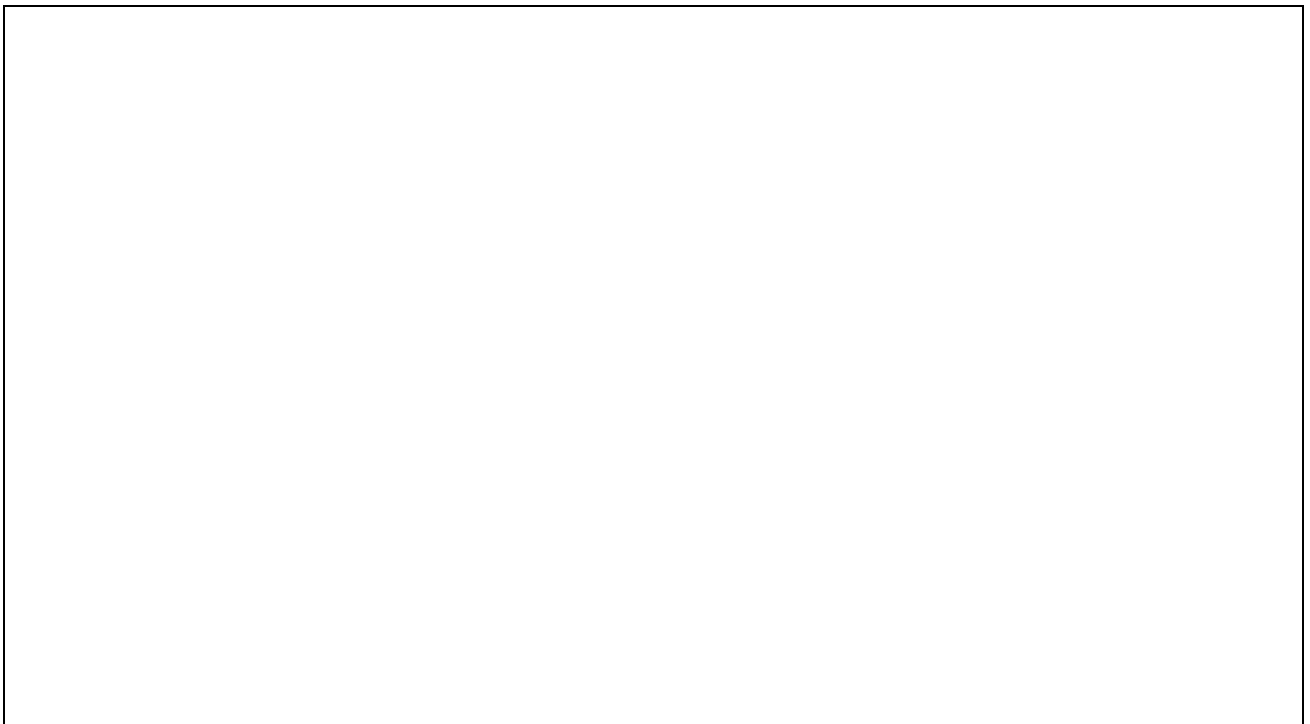
3. **[Softmax classifier]** Write your code to implement the softmax classifier. (Default hyperparameter settings for (b), (c): Epoch = 150, Batch_size = 512, learning_rate = 0.0005)

(a) **[30 pts]** Implement functions in 'models/SoftmaxClassifier.py'. (**'forward'**, **'compute_grad'**, **'_softmax'**, and **'eval'** respectively). Given a mini-batch data (X, Y), the error function for a mini-batch is defined as follows:

$$E(w) = -\frac{1}{|\mathcal{B}|} \sum_{(x_i, y_i) \in \mathcal{B}} y_i * \log(\hat{y}_i),$$

$$where\ \hat{y}_i = softmax(w^T x_i), |\mathcal{B}|\ \textbf{is the number of the mini} - \textbf{batch samples}.$$

Fill in your code here. You also have to submit your code to i-campus.
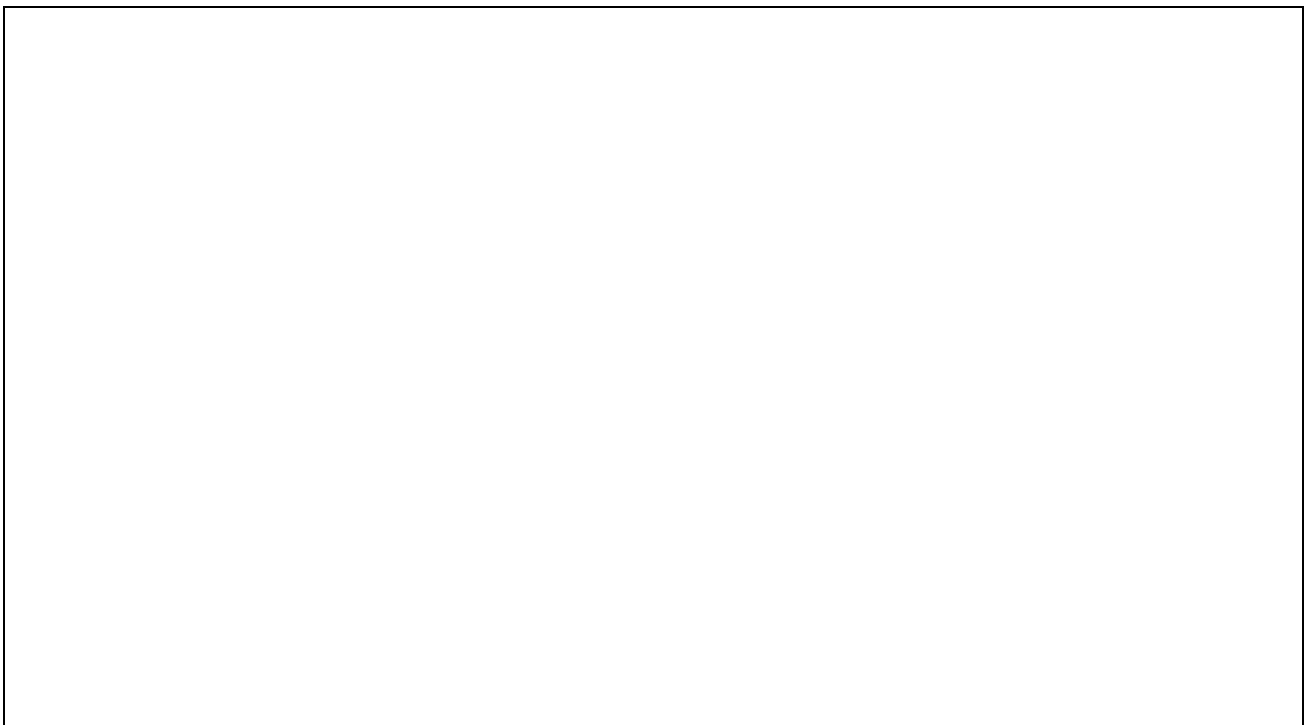
Answer:

(b) [5 pts] For the Litmus dataset, draw the plots by adjusting **the learning rate**. (The other hyperparameters are fixed as default hyperparameter settings.) The x-axis is the value of hyperparameters searched, and the y-axis is the accuracy score. Try at least five different values and explain your results.

Answer:

(c) [5 pts] For the Litmus dataset, draw the plots by adjusting the number of **epochs**. (The other hyperparameters are fixed as default hyperparameter settings.) The x-axis is the value of hyperparameters searched, and the y-axis is the accuracy score. Try at least five different values and explain your results.

Answer:

4. [BONUS CREDIT 20 pts] Write your review about the SKKU AI Colloquium held on Sep 22nd – 23rd in more than five lines. Please summarize the talk and write your comment about one or more lectures.

Answer: