**Fundamentals of Machine Learning (Fall 2022)**

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

**Student ID 2017314712**

**Name 김우진**

**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.
2. **[10 pts]** Suppose we have the following samples from the Bernoulli distribution. Calculate the parameter using the maximum likelihood estimation method.

{0, 1, 1, 1, 1, 0, 1, 1, 1, 0}

**Answer:**

|  |
| --- |
| **P( = 0, x2 = 1, x3=1, x4=1, x5=1, x6 = 0, x7 = 1, x8 = 1, x9 = 1, x10 = 0)**  **= (1-p)\*p\*p\*p\*p\*(1-p)\*p\*p\*p\*(1-p) -> 로그취하면**  **=ln(1-p)+ln(p)+ ln(p)+ ln(p)+ ln(p)+ ln(1-p)+ ln(p)+ ln(p)+ ln(p)+ ln(1-p) 미분**  **->**  **3p = 7 – 7p p = 7/10** |

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

{0.3, 0.5, 0.7, 0.8, 1.2}

**Answer:**

|  |
| --- |
| **F(x) =**  {\displaystyle {\frac {1}{\sigma {\sqrt {2\pi }}}}\;\exp \left(-{\frac {\left(x-\mu \right)^{2}}{2\sigma ^{2}}}\right)\!}    **=**  **제곱이니 0.092** |

**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](https://archive.ics.uci.edu/ml/datasets/banknote+authentication) |
| 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.

1. **[Logistic regression]** Write your code to implement logistic regression. (Default hyperparameter settings for (b), (c): Epoch = 50, Batch\_size = 512, learning\_rate = 0.1)
2. **[30 pts]** Implement functions in 'models/LogisticRegression.py'. ('**forward**', '**compute\_grad**', '**\_sigmoid**', and '**eval**' respectively). Given a mini-batch data the error function for a mini-batch is defined as follows:

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

**Answer:**

|  |
| --- |
|  |

1. **[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:**

|  |
| --- |
| Learning rate이 증가될수록 accuracy가 증가되는 모습을 보였다. |

1. **[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:**

|  |
| --- |
| **Epoch가 증가될수록 accuracy도 증가되었다.** |

1. **[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)
2. **[30 pts]** Implement functions in 'models/SoftmaxClassifier.py'. ('**forward**', '**compute\_grad**', '**\_softmax**', and '**eval**' respectively). Given a mini-batch data the error function for a mini-batch is defined as follows:

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

**Answer:**

|  |
| --- |
|  |

1. **[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:**

|  |
| --- |
| **Learning rate이 증가될수록 정확도도 높아졌다가 어느순간 오히려 낮아지는 모습을 보이고있다.** |

1. **[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:**

|  |
| --- |
| **어느 일정 epoch 까지는 정확도가 증가했지만 어느순간 기점으로부터는 정확도가 거의 변하지 않았다.** |

**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:**

|  |
| --- |
| **반 지도 학습(SSL)은 적은 수의 레이블이 지정된 데이터와 함께 쉽게 액세스할 수 있는 레이블이 없는 많은 데이터를 사용하여 더 나은 모델을 만드는 방법이다. 기존 SSL 연구의 대부분은 충분한 양의 레이블이 지정된 샘플, 수십에서 수백 개의 레이블이 지정된 샘플을 사용할 수 있는 경우에 중점을 둡니다. 하지만 여기서는 희귀 라벨 샘플, 1개 또는 2개의 라벨 샘플에**  **집중한다. 기존 접근 방식과 비교하여 8.9%에서 120.2%로 개선되었다. 많은 양에 데이터가 필요하지 않음에도 불구하고 높은 정확도를 보이면 데이터를 얻는 비용적인 측면에서 크게 줄일수 있기에 매우 큰 이득이라고 생각한다.** |