



作业 4 (PPT)

- 下表是3例近视样本数据，其标签依次为{0,1,2}，其中0代表正常样本，1代表轻度近视，2代表重度近视。表中给出了模型将样本预测为各个类的结果，请根据表中数据写出样本真实标签对应的one-hot编码并计算交叉熵损失。**log计算统一以e为底。**

编号	正常	轻度近视	重度近视
1	0.6	0.3	0.1
2	0.2	0.5	0.3
3	0.0	0.25	0.75



作业 4 (PPT)

1. Import data

```
[1] ✓ 0.3s

import numpy as np
import pandas as pd

df = pd.read_csv('HW4.csv')
df = df.drop('编号', axis=1)
df
```

	正常	轻度近视	重度近视
0	0.6	0.30	0.10
1	0.2	0.50	0.30
2	0.0	0.25	0.75

2. Use one-hot code for real label

```
[3] ✓ 0.0s Python

real_onehot = [[1,0,0], [0,1,0], [0,0,1]]
real_onehot = np.array(real_onehot)
real_onehot
```

```
... array([[1, 0, 0],
          [0, 1, 0],
          [0, 0, 1]])
```

One-hot code

```
[4] ✓ 0.0s

predict = np.array(df)
predict
```

```
... array([[0.6 , 0.3 , 0.1 ],
          [0.2 , 0.5 , 0.3 ],
          [0. , 0.25, 0.75]])
```

3. Compute the cross-entropy

```
[5] ✓ 0.0s Python

# compute the cross entropy
def cross_entropy(real_onehot, predict):
    return -np.sum(real_onehot*np.log(predict+1e-8))

cross_entropy(real_onehot, predict)
```

```
... np.float64(1.4916548267777172)
```

Cross-entropy