Machine Learning: Project 2

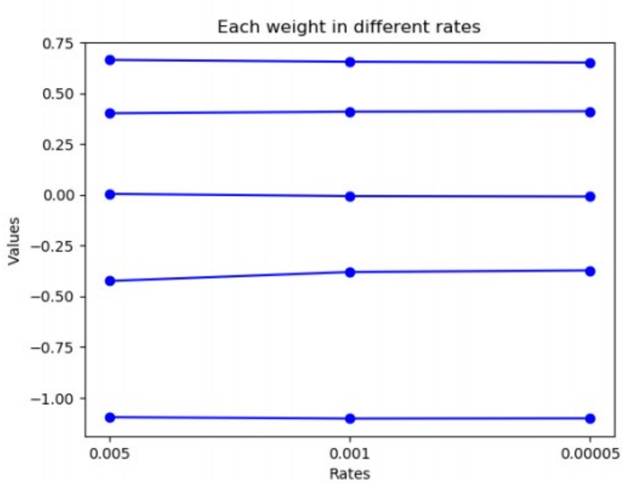
Liam Li G48582460

Head:

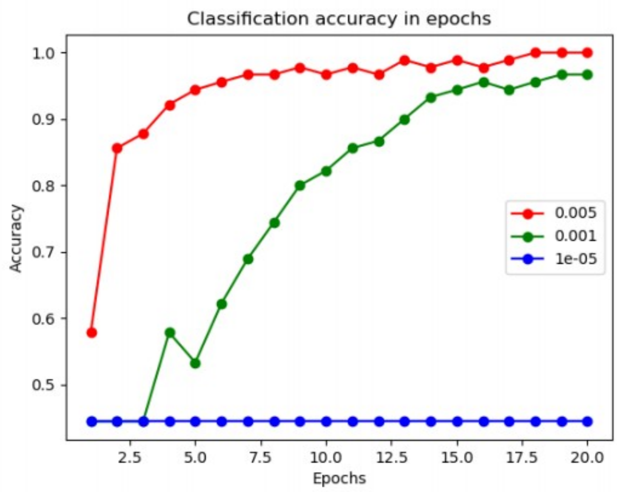
Python: pandas and matplotlib modules for visualization
Sublime text 3

1. When the learning rate increases from 0.00005 to 0.001 and to 0.005, the accuracy doesn't change, but the time of training epochs is getting bigger, and some weights are slightly changed.

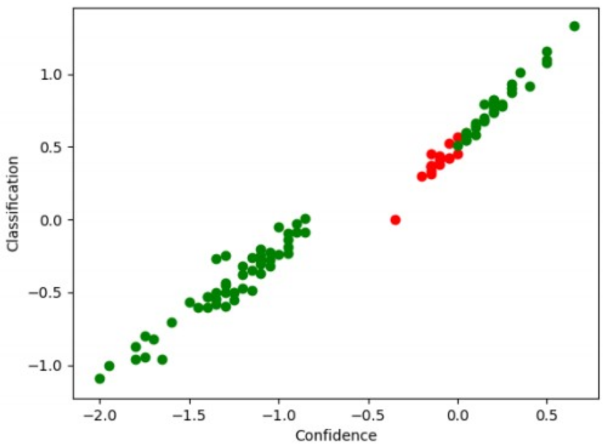
```
weights:  [-1.095, 0.005, 0.665, -0.424, 0.402]
learning rate:  0.005
accuracy:  1.0
weights:  [-1.102, -0.006, 0.656, -0.38, 0.41]
learning rate:  0.001
accuracy:  1.0
weights:  [-1.101, -0.008, 0.652, -0.372, 0.412]
learning rate:  1e-05
accuracy:  1.0
[Finished in 3.2s]
```



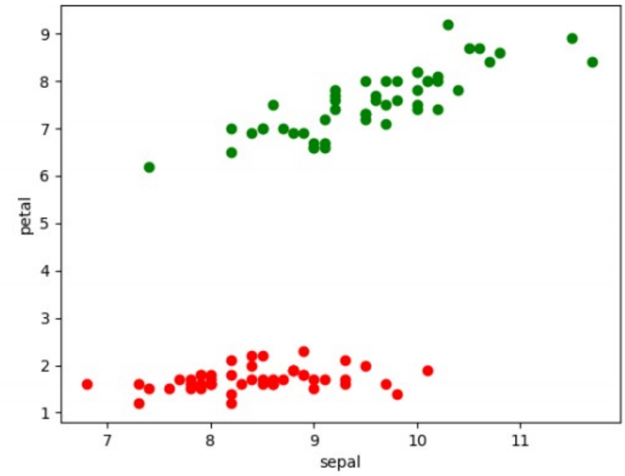
2. It does converge otherwise it will never stop iteration to get



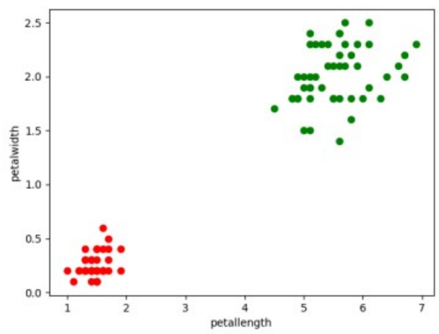
3. When I use wights[-1,0,0.5,-0.5,0.5] as final weights to execute activation function, the confidence is 83%.
   The red points below are misclassification points.

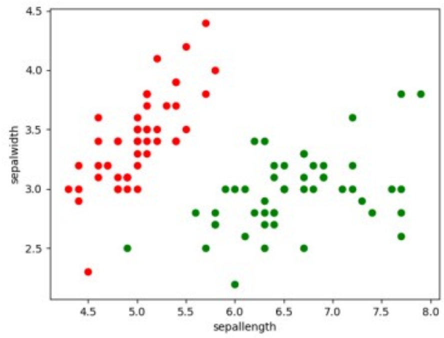


4. Yes, it is linearly separable.



I create this plot by summing up the length and width of sepal and petal, because all four parameters affect the result.





1. $\log L(\lambda) = \log P(X|\lambda) = \sum_{i=1}^{n} \log(X_i|\lambda)$

$$= \sum_{i=1}^{n} (X_i \cdot \log\lambda - \log X_i!) - n\lambda$$

$$\frac{d}{dx} = \log L(\lambda) = \frac{\sum_{i=1}^{n} X_i}{\lambda} - n = 0$$

$$\hat{\lambda} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$$E[\hat{\lambda}] = E\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\sum_{i=1}^{n} E[X_i] = \frac{1}{n}\sum_{i=1}^{n} \lambda = \lambda$$

2. $P(\lambda|y) \propto \left(\frac{\beta^{\alpha}}{\Gamma(\alpha)} \cdot \lambda^{\alpha-1} \cdot e^{-\beta\lambda}\right) \cdot \left(\frac{e^{-n\lambda}\lambda^{\sum_{i=1}^{n} X_i}}{\prod_{i=1}^{n} X_i!}\right)$

$$= \frac{\beta^{\alpha}}{\Gamma(\alpha)\prod_{i=1}^{n} X_i!} \cdot \left(\lambda^{\alpha-1} + \sum_{i=1}^{n} X_i e^{-(n+\beta)\lambda}\right)$$

$$P(\lambda|y) \propto \lambda^{\hat{\alpha}-1} e^{-\hat{\beta}\lambda} = (\lambda|\alpha + \sum_{i=1}^{n} X_i, \beta + N)$$
$$= \Gamma(\alpha, \beta)$$

3. $\frac{\alpha - 1}{\beta} = \frac{\sum_{i=1}^{n} X_i + \alpha - 1}{n + \beta}$