HW2-Copy1

February 26, 2017

** Machine Learning HW2   Amirhossein Imani - ai2335 **

0.1 Problem 1 - a

To derive $\hat{\pi}$ from the objection function, we need to take the derivative of objective function based on π. It will be equal to:

$$\frac{\sigma}{\sigma\pi}\sum_{i=0}^{n} lnp(y_i|\pi) + \frac{\sigma}{\sigma\pi}\sum_{i=0}^{n} lnp(x_i1|\theta_{y_i}^{(1)}) + \frac{\sigma}{\sigma\pi}\sum_{i=0}^{n} lnp(x_i2|\theta_{y_i}^{(2)})$$

$$= \frac{\sigma}{\sigma\pi}\sum_{i=0}^{n} lnp(y_i|\pi) + 0 + 0 = \frac{\sigma}{\sigma\pi}\sum_{i=0}^{n} ln[\pi_i^x(1-\pi)^{1-x_i}]$$

$$= \frac{\sigma}{\sigma\pi}\sum_{i=0}^{n}[xln\pi + (1-x)ln(1-\pi)] = \frac{\sum_{i=0}^{n} x}{\pi} + \frac{\sum_{i=0}^{n} 1-x}{1-\pi}$$

$$\frac{\sum_{i=0}^{n} x}{\pi} - \frac{\sum_{i=0}^{n} 1-x}{1-\pi} = 0 => \frac{\sum_{i=0}^{n} x}{\pi} = \frac{n-\sum_{i=0}^{n} x}{1-\pi} =>$$

$$\hat{\pi}_{mle} = \frac{\sum_{i=0}^{n} x}{n}$$

0.2 Problem 1 - b

Steps are like above. first we have to take the derivative in respect to $\hat{\theta}_y^{(1)}$, equal that to zero and solve it.

$$\frac{\sigma}{\sigma{\theta_y}^{(1)}}\sum_{i=0}^{n} lnp(y_i|\pi) + \frac{\sigma}{\sigma{\theta_y}^{(1)}}\sum_{i=0}^{n} lnp(x_{i1}|\theta_{y_i}^{(1)}) + \frac{\sigma}{\sigma{\theta_y}^{(1)}}\sum_{i=0}^{n} lnp(x_{i2}|\theta_{y_i}^{(1)})$$

$$= 0 + \frac{\sigma}{\sigma{\theta_y}^{(1)}}\sum_{i=0}^{n} lnp(x_i1|\theta_{y_i}^{(1)}) + 0 = \frac{\sigma}{\sigma\theta_{y_i}^{(1)}}\sum_{i=0}^{n} ln[\theta_i^x(1-\theta_{y_i}^{(1)})^{1-x_i}] = \frac{\sum_{i=0}^{n} x}{\theta_{y_i}^{(1)}} + \frac{\sum_{i=0}^{n} 1-x}{1-\theta_{y_i}^{(1)}}$$

$$\frac{\sum_{i=0}^{n} x}{\theta_{y_i}^{(1)}} - \frac{\sum_{i=0}^{n} 1-x}{1-\theta_{y_i}^{(1)}} = 0 => \frac{\sum_{i=0}^{n} x}{\theta_{y_i}^{(1)}} = \frac{n-\sum_{i=0}^{n} x}{1-\theta_{y_i}^{(1)}} =>$$

$$\hat{\theta_{y_i}^{(1)}}_{mle} = \frac{\sum_{i=0}^{n} x}{n}$$

0.3 Problem 1 - c

Steps are like above. first we have to take the derivative in respect to $\hat{\theta}_y^{(2)}$, equal that to zero and solve it.

$$\frac{\sigma}{\sigma\theta_y{}^{(2)}}\sum_{i=0}^{n} lnp(y_i|\pi) + \frac{\sigma}{\sigma\theta_y{}^{(2)}}\sum_{i=0}^{n} lnp(x_{i1}|\theta_{y_i}^{(1)}) + \frac{\sigma}{\sigma\theta_y{}^{(2)}}\sum_{i=0}^{n} lnp(x_{i2}|\theta_{y_i}^{(2)})$$

$$= 0+0+\frac{\sigma}{\sigma\theta_y{}^{(2)}}\sum_{i=0}^{n} lnp(x_{i2}|\theta_{y_i}^{(2)}) = \frac{\sigma}{\sigma\theta_{y_i}^{(2)}}\sum_{i=0}^{n} ln[\theta_y^{(2)}(x_{0,2})^{-(\theta_y^{(2)}+1)}] = \frac{\sigma}{\sigma\theta_{y_i}^{(2)}}\sum_{i=0}^{n}[ln\theta_y^{(2)} - (\theta_y^{(2)}+1)ln(x_{0,2})]$$

$$\frac{\sigma}{\sigma\theta_{y_i}^{(2)}}\sum_{i=0}^{n}[ln\theta_y^{(2)} - (\theta_y^{(2)}+1)ln(x_{0,2})] = 0 => \frac{n}{\theta_{y_i}^{(2)}} - \sum_{i=0}^{n} ln(x_{0,2}) = 0 => \frac{n}{\theta_{y_i}^{(2)}} = \sum_{i=0}^{n} ln(x_{0,2}) =>$$

$$\hat{\theta_{y_i}^{(2)}}{}_{mle} = \frac{n}{\sum_{i=0}^{n} ln(x_{0,2})}$$

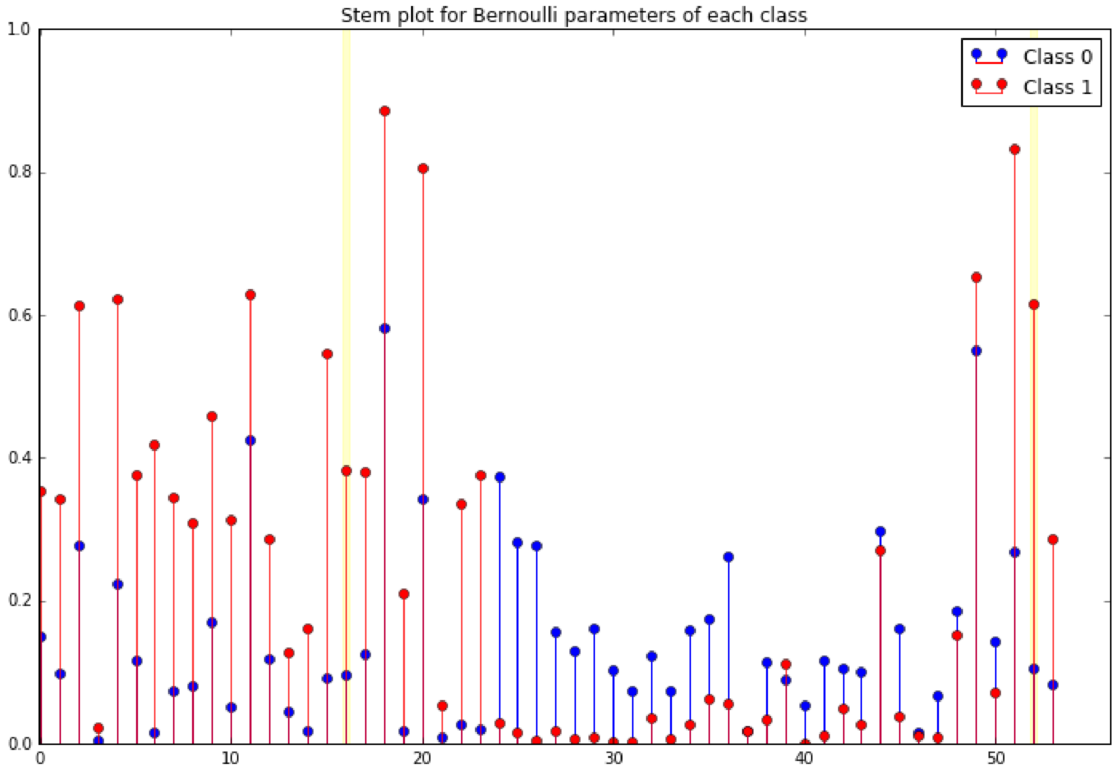
1 Problem 2

```
In [6]: print(df_confusion, '\n', '\n',"Accuracy:",accu)
```

```
Predicted   0   1
Actual
0.0        54   2
1.0         5  32

 Accuracy: 0.9247311827956989
time: 21.4 ms
```

1.1 B - Stem Plot

```
In [18]: plt.show()
```



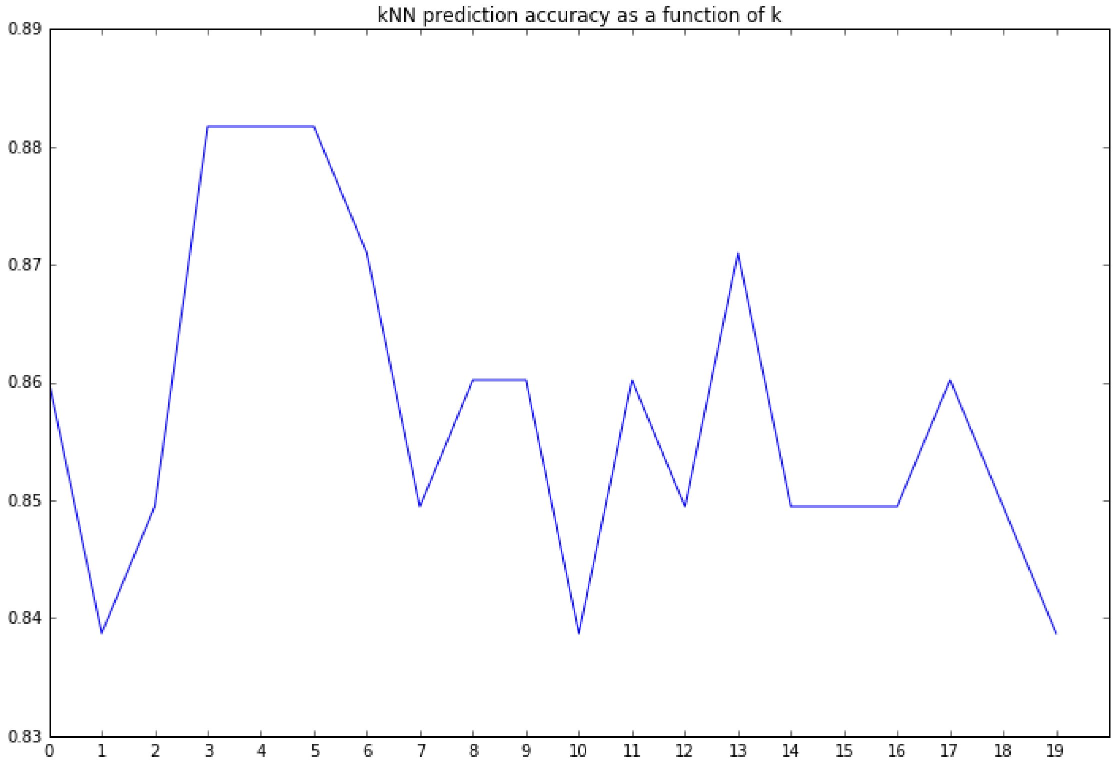

```
time: 665 ms
```

1.1.1 make observation about dimension 16 and 52

Those two words are “free” and “!” from the document description. Based on parameters stem plot, we can claim that the probability of an email being marked as spam if it contains those two words, is much higher than being not-spam.

1.2 C - KNN Algorithm

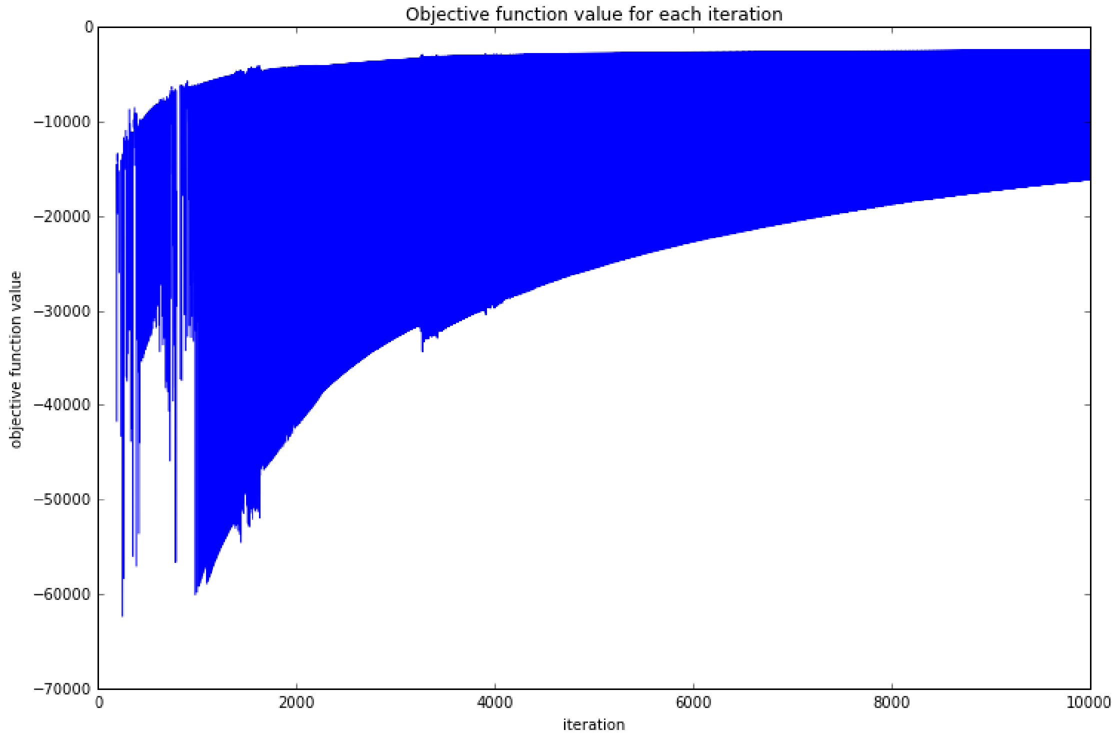
```
In [9]: # Chart title
        plt.title('kNN prediction accuracy as a function of k')

Out[9]: <matplotlib.text.Text at 0x1032058d0>
```




```
time: 2.4 s
```

2 D- Logistic Regression - steepest ascent algorithm

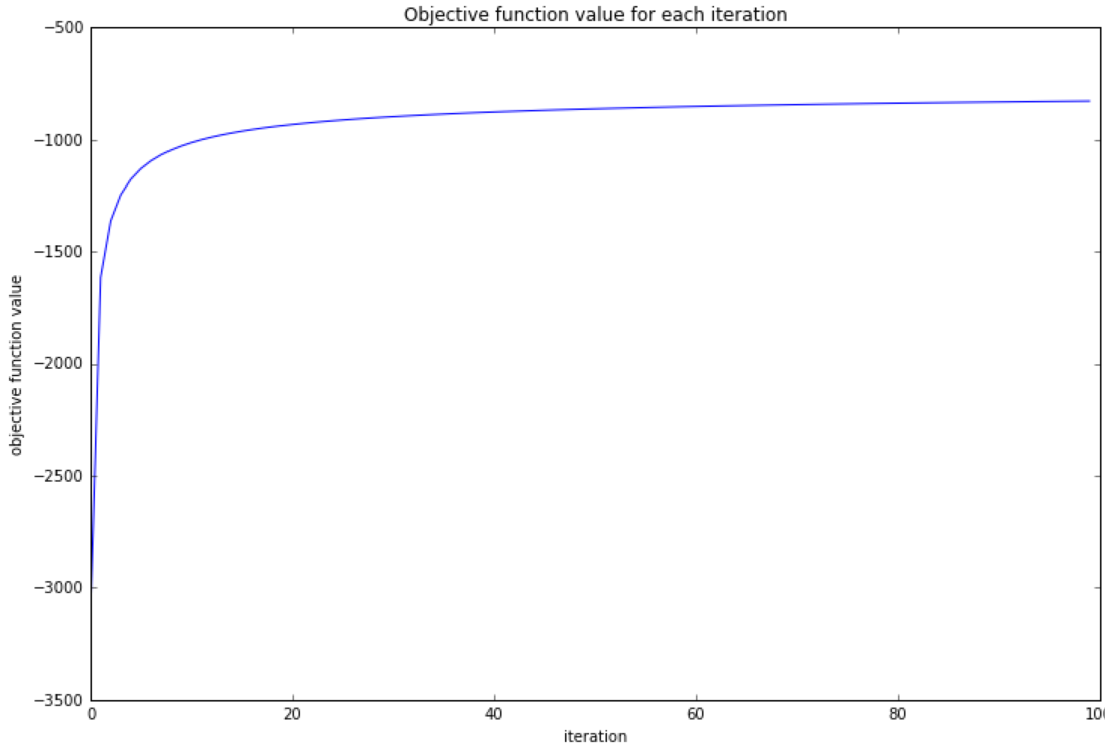
```
In [24]: plt.title('Objective function value for each iteration')

Out[24]: <matplotlib.text.Text at 0x11181c1d0>
```




```
time: 1.09 s
```

2.1 E- Newton Method

```
In [166]: figtext(.95, .9, "Accuracy: 0.913978494623655 88")
```


Objective function value for each iteration
objective function value
iteration
−500
−1000
−1500
−2000
−2500
−3000
−3500
0
20
40
60
80
100

`time: 451 ms`