(a) What was the estimated value of P(y) for y = 1?

```
0.13448193347690834
```

(b) What was the estimated value of P(y) for y = 0?

```
0.8655180665230916
```

(c) What were the estimated values for $\phi_{admirer|y}$ for the corresponding feature admirer when y = 1 and for y = 0.

```
The estimated value of phi_admirer for y = 1 is:  0.014234875444839
857
The estimated value of phi_admirer for y = 0 is:  0.0
```

(d) What were the estimated values for $\phi_{secret|y}$ for the corresponding to feature secret when y = 1 and for y = 0.

```
The estimated value of phi_secret for y = 1 is:  0.0142348754448398
57
The estimated value of phi_secret for y = 0 is:  0.0005529444290848
77
```

(e) Which classes were predicted for the first 5 examples in the test set?

```
[0, 0, 0, 0, 0]
```

(f) Which classes were predicted for the last 5 examples in the test set?

```
[0, 1, 1, 0, 0]
```

(g) What was the percentage error on the examples in the test file?

```
0.048815506101938265
```

(h) Repeat the above step (question 4g) by adding m smoothing by trying different values of m where 0 < m <= 1. Did the smoothing help? If so, for what value of m?

Yes, the smoothing can decrease the error percentage.

```
When m = 0.05, the error rate is minimum: 0.013639626704953339
```

(i) What accuracy is attained is you use Zero-R instead of Bernoulli Naive Bayes?

```
The label which Zero-R predicts is:  0
The error rate of Zero-R is:  0.13280689160086145
```