Homework 2

Yuan Li, N19728558, NetID: yl6606

Part II: Programming and Questions

(a) Answer:

The estimated value of $P(y)$ for $y = 1$: $P(y_1) = 0.13448$.

(b) Answer:

The estimated value of $P(y)$ for $y = 0$: $P(y_0) = 0.86552$.

(c) Answer:

The estimated values for $\phi_{admirer}|y$ for the corresponding feature **admirer** when $y = 1$ and for $y = 0$.

$$\phi_{admirer}|y_0 = 0.0$$
$$\phi_{admirer}|y_1 = 0.0016750$$

(d) Answer:

The estimated values for $\phi_{secret}|y$ for the corresponding feature **secret** when $y = 1$ and for $y = 0$.

$$\phi_{secret}|y_0 = 0.0009572$$
$$\phi_{secret}|y_1 = 0.0019143$$

(e) Answer:

Classes for the first 5 examples in the test set: [ham, ham, ham, ham, ham]

(f) Answer:

Classes for the last 5 examples in the test set: [ham, ham, ham, ham, ham]

Note: All the answers above are based on $m = 0$, which means without smoothing.

(g) Answer:

The percentage error on the examples in the test file: $P(error) = 13.28\%$

(h) Answer:

(i) Answer:

The accuracy on the test examples using zero-R: $acc = 86.72\%$