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 $y = 0$.

$$\phi_{admirer}|y_0 = 0.0$$
$$\phi_{admirer}|y_1 = 0.0142349$$

(d) Answer:

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

$$\phi_{secret}|y_0 = 0.0005529$$
$$\phi_{secret}|y_1 = 0.0142349$$

(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, spam, spam, ham, ham]

(g) Answer:

The percentage error on the examples in the test file: $P(error) = 0.0488155$

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

(h) Answer:

I tested $m = 0.1 * i, i = 1, 2, ..., 10$, and the error percentage is as follow:

$$error_{m=0} = 0.0488155$$
$$error_{m=0.1} = 0.0136396$$
$$error_{m=0.2} = 0.0143575$$
$$error_{m=0.3} = 0.0157932$$
$$error_{m=0.4} = 0.0157932$$
$$error_{m=0.5} = 0.0150754$$
$$error_{m=0.6} = 0.0157932$$
$$error_{m=0.7} = 0.0165111$$
$$error_{m=0.8} = 0.0208183$$
$$error_{m=0.9} = 0.0222541$$
$$error_{m=1.0} = 0.0222541$$

If we add m, the error percentage becomes smaller, and it is smallest when $m = 0.1$. So the smoothing does help improve the accuracy, the value of $m = 0.1$.

(i) Answer:

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