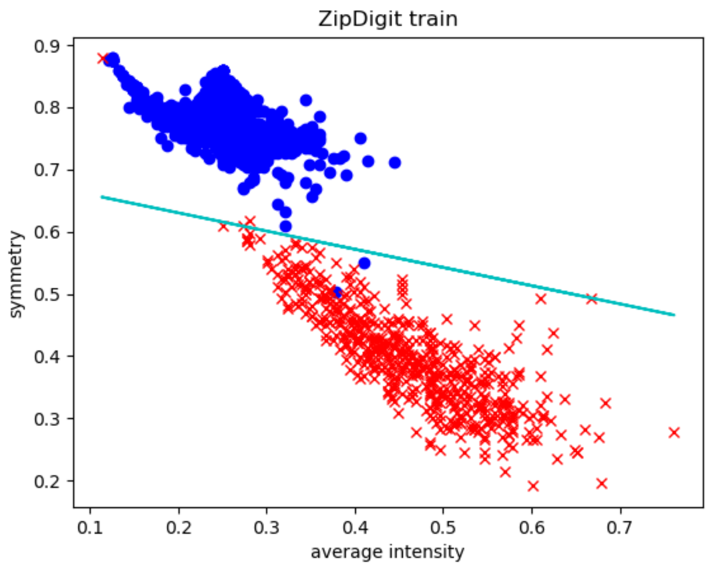
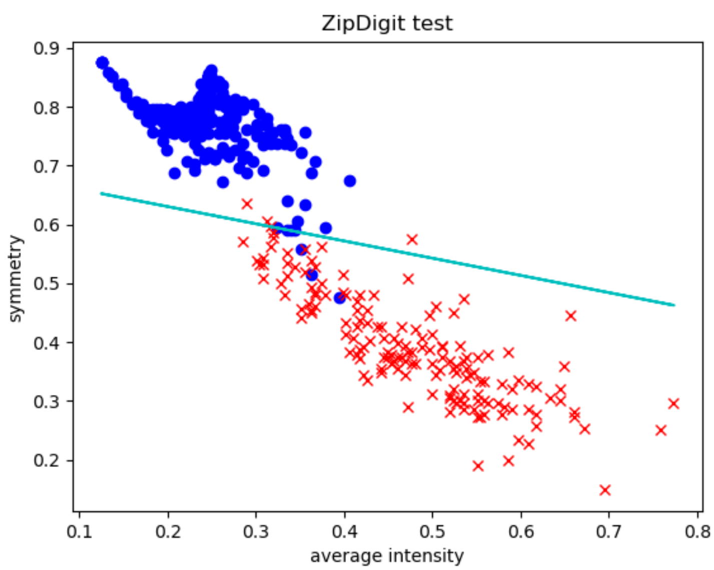
1.

a)





b)

```
E_in is: 0.081
E_test is: 0.127
```

$E_{in} \approx 0.081$

$E_{test} \approx 0.127$

c)

Apply $\Omega(N, H, \delta) = \sqrt{\frac{8}{N} In(\frac{4m_H(2N)}{\delta})} \le \sqrt{\frac{8}{N} In(\frac{4((2N)^{d_{vc}}+1)}{\delta})}$

In linear classification d_{vc} = d + 1

Training set:

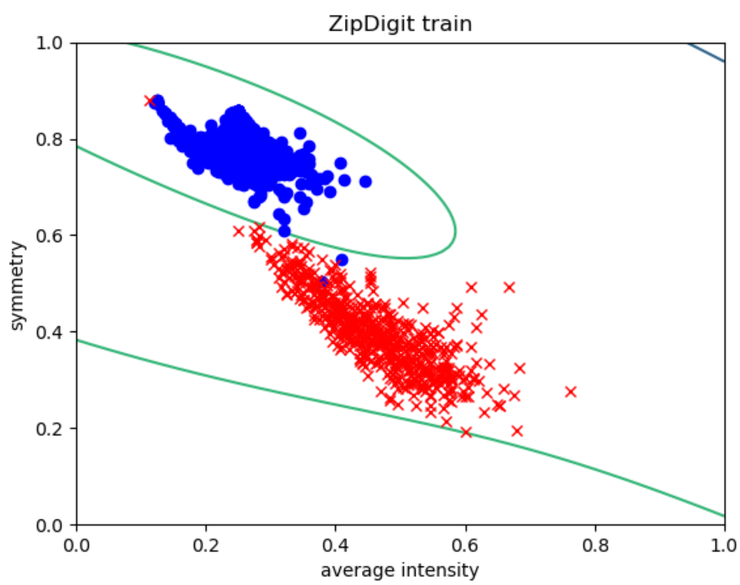
$$\sqrt{\frac{8}{1561} In(\frac{4((2\times1561)^3+1)}{0.05})} \approx 0.382$$

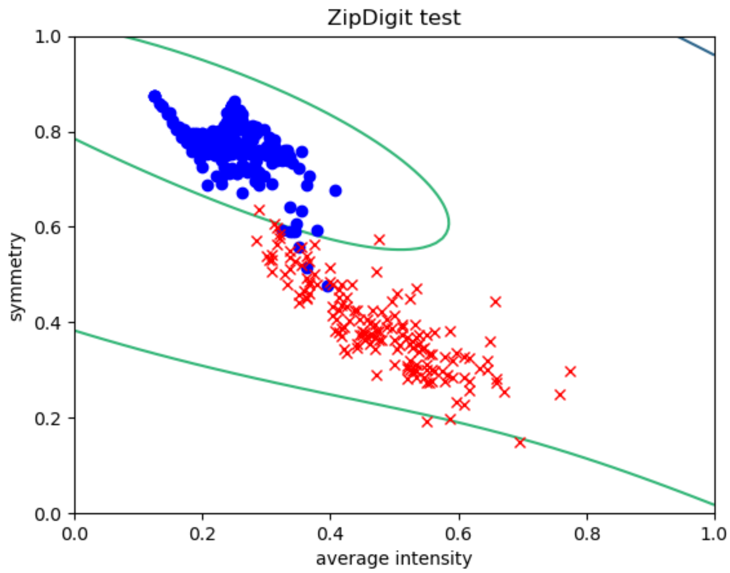
Test set:

$$\sqrt{\frac{8}{424} In(\frac{4((2\times424)^3+1)}{0.05})} \approx 0.681$$

The bound for the training set is better.

d)





For 3rd order polynomial transform,

```
E_in is: 0.030
E_test is: 0.083
```

$E_{in} \approx 0.030$

$E_{test} \approx 0.083$

Apply $\Omega(N, H, \delta) = \sqrt{\frac{8}{N} In(\frac{4m_H(2N)}{\delta})} \leq \sqrt{\frac{8}{N} In(\frac{4((2N)^{d_{vc}}+1)}{\delta})}$

In 3rd order polynomial transform, the d_{vc} = 9 + 1 = 10

Training set:

$$\sqrt{\frac{8}{1561} In(\frac{4((2\times1561)^{10}+1)}{0.05})} \approx 0.659$$

Test set:

$$\sqrt{\frac{8}{424} In(\frac{4((2\times424)^{10}+1)}{0.05})} \approx 1.164$$

The bound for the training set is better.

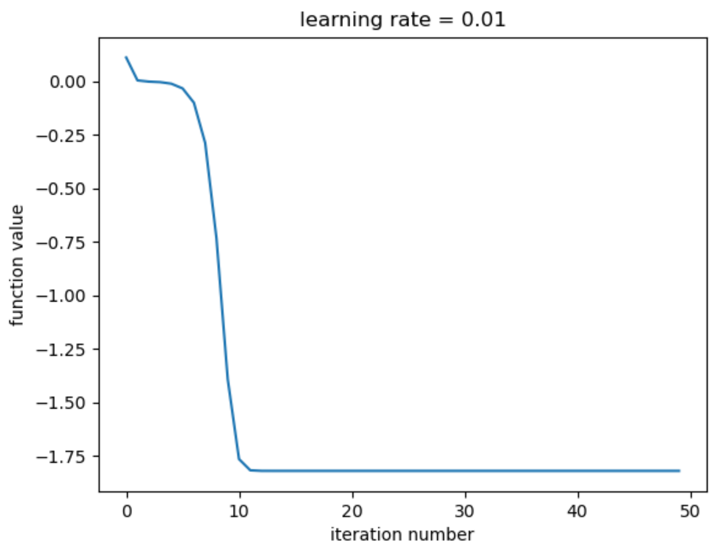
e)

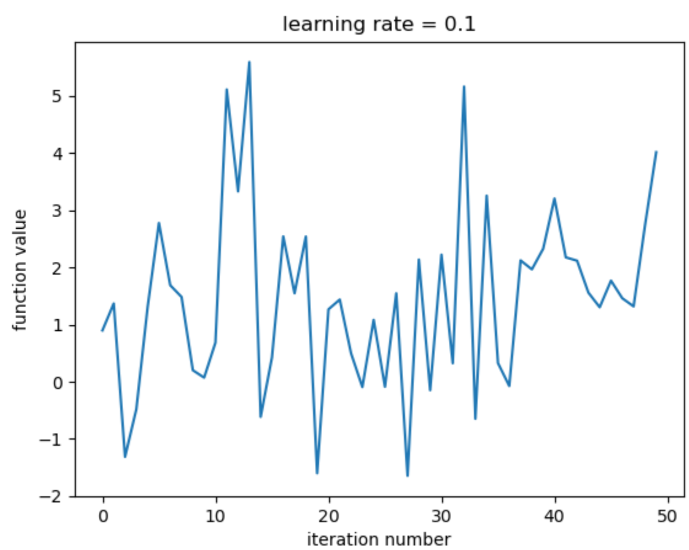
I will choose the 3rd order polynomial transform to the customer.

As computed in d) and b), the E_{in} and E_{test}(stands for E_{out})has been improved a lot from standard feature space to 3rd order polynomial feature space. This shows a better in-sample fit and out-of-sample fit.

2.

a)





When changing from 0.01 to 0.1, the value bounced around when approaching the minimum due to the large learning rate, which differs from learning rate 0.01 as it gradually approaches minimum and coverage.

b)

Start location	(0.1, 0.1)	(1, 1)	(−0.5, −0.5)	,(−1, −1)

Min value	-1.820	0.593	-1.332	0.593

Problem 3.16

a)

Since $g(x) = P[y = +1 | x]$, we can get that $P[y = -1 | x] = 1 - g(x)$.

Hence, cost(accept) = $g(x) \times 0 + (1 - g(x)) \times c_a = (1 - g(x)) \times c_a$

cost(reject) = $g(x) \times c_r + (1 - g(x)) \times 0 = g(x) \times c_r$

b)

From a) we have

cost(accept) = $(1 - g(x)) \times c_a$

cost(reject) = $g(x) \times c_r$

We only accept when cost(accept) <= cost(reject),

$(1 - g(x)) \times c_a \leq g(x) \times c_r$

$c_a \leq g(x)(c_a + c_r)$

$\frac{c_a}{c_a + c_r} \leq g(x)$

According to the problem, we have,

$\kappa = \frac{c_a}{c_a + c_r}$

c)

Supermarket:

$\kappa = \frac{1}{1+10} = \frac{1}{11}$

CIA:

$\kappa = \frac{1000}{1000+1} = \frac{1000}{1001}$