1a) $Y = \beta_1 X_1 + \beta_2 X_2 + \beta_0$

1b) β_0 is avg. lowering of cholesterol when $X_1 = X_2 = 0$ (female, no drug)
β_1 is avg lowering of cholesterol for both genders when drug is used
β_2 is avg. difference in cholesterol b/w men & women

1c) We can determine significance by using a t-test. We use our model to fit n datasets, & generate means & std dev for β_2. We then calculate the t-test, & use p-values to determine significance. For ex, threshold of 0.05, so if p-value < 0.05 we reject that effect of drug was significantly diff in men compared to women

2a) (i) On avg, someone who did not receive treatment & doesn't prefer A is 3.135 times more likely to prefer A the second time they're asked

ii) $\log\left(\frac{p}{1-p}\right) = 3.135$ $\frac{p}{1-p} = e^{3.135}$

iii) yes

2b) (i) On avg. someone who received the treatment is 2.309 times less likely to prefer A the second time they're asked

(ii) 0.433

(iii) A is far less popular among those who received the treatment

2c) (i) On avg., someone who preferred A initially is 5.15 times less likely to prefer A the second time they're asked

(ii) 5.15

(iii) This implies that the treatment increases the odds of someone preferring A to B

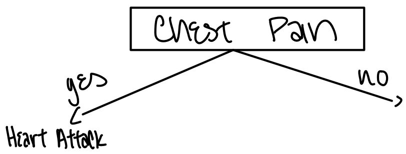
2d) (i) On avg, a one unit increase in Treatment × Prefer A will result in a 2.85 greater chance of an individual preferring A the second time they're asked.

3a)

Feature	Gini
Chest pain	0.222
male	0.333
smokes	0.417
exercises	0.333

CP&HA : 3/3 , CP&!HA: 0/3 => 0
!CP&HA : 1/3, !CP&!HA: 2/3 => 0.44
M&HA : 2/4 , M&!HA: 2/4 => 0.5
!M&HA: 2/2, !M&!HA: 0/2 => 0
S&HA : 3/4 , S&!HA: 1/4 => 0.375
!S&HA: 1/2, !S&!HA: 1/2 => 0.5
E&HA : 2/4 , E&!HA: 2/4 => 0.5
!E&HA : 2/2 , !E&!HA: 0/2 => 0

Chest pain has lowest Gini index => split

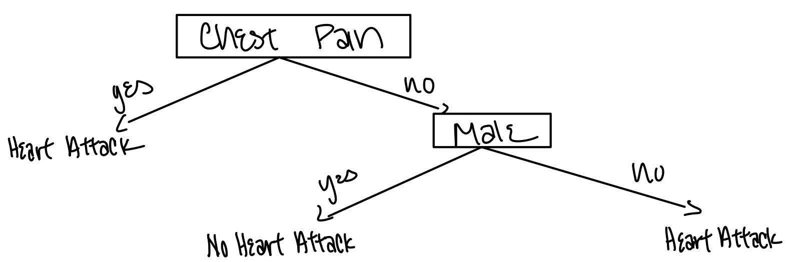


PATIENT ID	CHEST PAIN?	MALE?	SMOKES?	EXERCISES?	HEART ATTACK?
1.	yes	yes	no	yes	yes
2.	yes	yes	yes	no	yes
3.	no	no	yes	no	yes
4.	no	yes	no	yes	no
5.	yes	no	yes	yes	yes
6.	no	yes	yes	yes	no

Feature	Gini
Male	0
smokes	0.33
exercises	0

M&HA : 0/2 , M&!HA: 2/2 => 0
!M&HA: 1/1 , !M&!HA: 0/1 => 0
S&HA : 1/2 , S&!HA : 1/2 => 0.5
!S&HA: 0/1 , !S&!HA: 1/1 => 0
E&HA : 0/2 , E&!HA: 2/2 => 0
!E&HA : 1/1 , !E&!HA: 0/1 => 0

Male & exercise have the same gini index, so we split on male



3b) CP → HA

!CP & M → !HA

!CP & !M → HA

4a) Hidden 1 $= (I_1)(W11) + (I_2)(W21) + 0.1$

$= (0.1)(1) + (0.2)(0.1) + 0.1 = 0.22 \Rightarrow \frac{1}{1+e^{-0.22}} = 0.555$

Hidden 2 $= (I_1)(W12) + (I_2)(W22) + 0.1$

$= (0.1)(0.5) + (0.2)(0.2) + 0.1 = 0.19 \Rightarrow \frac{1}{1+e^{-0.19}} = 0.547$

Output $= (W_{13})(net_{h_1}) + (W_{23})(net_{h_2}) + 0.5$

$(1)(0.555) + (0.5)(0.547) + 0.5 = 1.328 \Rightarrow \frac{1}{1+e^{-1.328}} = 0.791$

$y = 1,\ \hat{y} = 0.791$

Loss $= -(y\log(\hat{y}) + (1-y)\log(1-\hat{y})) = 0.235$

4b) $\frac{\partial loss}{\partial W_{11}} = \frac{\partial loss}{\partial out} \cdot \frac{\partial out}{\partial H_1} \cdot \frac{\partial H_1}{\partial W_{11}}$

Loss $= -(y\log(\hat{y}) + (1-y)\log(1-\hat{y}))$

$\frac{\partial loss}{\partial out} = -\left(\frac{y}{\hat{y}} - \frac{1-y}{1-\hat{y}}\right) = -\left(\frac{1}{0.791} - \frac{1-1}{0.791}\right)$

$= -1.264$

$Out = \frac{1}{1+e^{-(W_{13}H_1 + W_{23}H_2 + 0.5)}}$

$\frac{\partial out}{\partial H_1} = -(1+e^{-(W_{13}H_1 + W_{23}H_2 + 0.5)})^{-2}(-W_{13}e^{-(W_{13}H_1 + W_{23}H_2 + 0.5)})$

$= 0.1656$

$H_1 = \frac{1}{1+e^{-(W_{11}I_1 + W_{21}I_2 + 0.1)}}$

$\frac{\partial H_1}{\partial W_{11}} = -(1+e^{-(W_{11}I_1 + W_{21}I_2 + 0.1)})^{-2}(-I_1 e^{-(W_{11}I_1 + W_{21}I_2 + 0.1)})$

$= 0.0247$

$\frac{\partial loss}{\partial out} = (-1.264)(0.1656)(0.0247) = -0.0052$

4c) $W11' = W11 - \mu\left(\frac{\partial loss}{\partial W11}\right)$

$= 1 - (0.1)(-0.0052) = 1.00052$

Hidden 1' $= (I_1)(W11') + (I_2)(W21) + 0.1$

$= (0.1)(1.0005) + (0.2)(0.1) + 0.1 = 0.22 \Rightarrow \frac{1}{1+e^{-0.22}} = 0.555$

Hidden 2' $= (I_1)(W12) + (I_2)(W22) + 0.1$

$= (0.1)(0.5) + (0.2)(0.2) + 0.1 = 0.19 \Rightarrow \frac{1}{1+e^{-0.19}} = 0.547$

Output' $= (W_{13})(net_{h_1}) + (W_{23})(net_{h_2}) + 0.5$

$(1)(0.555) + (0.5)(0.547) + 0.5 = 1.328 \Rightarrow \frac{1}{1+e^{-1.328}} = 0.7906$

$y = 1,\ \hat{y} = 0.7906$

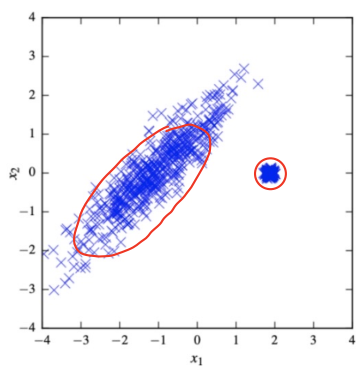
Loss $= -(\log(0.7906)) = 0.235$

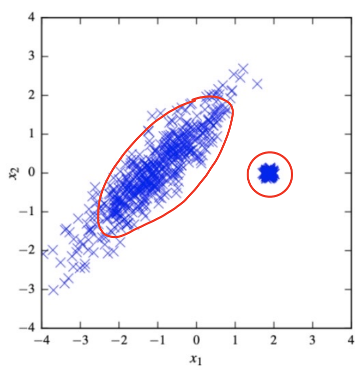
5a) $K=2$, $\mu^k = c_k$

Since $\|\bar{x}_k - c_k\| \le d$, we know that after the first itr., we won't move more than d, so no new pts. added to cluster, so it converges

5b) if d large enough, we will never re-cluster & gain / lose points in cluster, so converge after first itr

5d) DBSCAN b/c density based





6.1) Linear SVM - (c): linear, equidistant from both light

6.2) Kernelized SVM (order 2 poly): (d)

6.3) Perceptron - (b) - its linear, but has slope so not logistic regression or SVM

6.4) Logistic regression - (a): linear lower boundary b/c far away from red

6.5) Decision tree: (g) - very precise, overfitted

6.6) Random Forest: (h) - similar to decision but less precise b/c random sample of training data

6.7) Neural network (10 ReLU): (f): NN, but rough lines

6.8) Neural Network (10 tanh units) - (e): NN, smoother lines